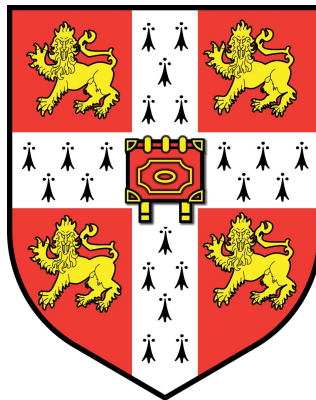


Geometric Aspects of Gauge and Spacetime Symmetries

Steffen Christian Martin Gielen
Trinity College

A dissertation submitted to the
University of Cambridge
for the degree of
Doctor of Philosophy



January 2011

Declaration

This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text (section 1.2). It is not substantially the same as any other thesis that I have submitted or will submit for a degree at this or any other university.

This thesis is based on the research presented in the following refereed publications:

- (1) GARY W. GIBBONS AND STEFFEN GIELEN. *The Petrov and Kaigorodov-Ozsváth Solutions: Spacetime as a Group Manifold*. Classical and Quantum Gravity **25** (2008), 165009.
- (2) GARY W. GIBBONS AND STEFFEN GIELEN. *Deformed General Relativity and Torsion*. Classical and Quantum Gravity **26** (2009), 135005.
- (3) GARY W. GIBBONS, STEFFEN GIELEN, C. N. POPE, AND NEIL TUROK. *Measures on Mixing Angles*. Physical Review D **79** (2009), 013009.
- (4) STEFFEN GIELEN. *CP Violation Makes Left-Right Symmetric Extensions With Non-Hermitian Mass Matrices Appear Unnatural*. Physical Review D **81** (2010), 076003.
- (5) STEFFEN GIELEN AND DANIELE ORITI. *Classical general relativity as BF-Plebanski theory with linear constraints*. Classical and Quantum Gravity **27** (2010), 185017.

These papers appear in the text as references [65], [66], [67], [68], and [69] respectively. The results of section 5 of (3) were also separately published in

- (6) GARY W. GIBBONS, STEFFEN GIELEN, C. N. POPE, AND NEIL TUROK. *Naturalness of CP Violation in the Standard Model*. Physical Review Letters **102** (2009), 121802.

Steffen Gielen
Sunday, 13 January 2013

Acknowledgments

My first and foremost thanks go to my supervisor, Prof. Gary Gibbons, for teaching me so much about physics and mathematics, for many fruitful discussions and suggestions, for his encouragement and for his support of all of my plans over the last three years.

I have also greatly enjoyed working with, and learning from, Dr. Daniele Oriti, Prof. C. N. Pope and Dr. Neil Turok on different projects that are part of this thesis. Invitations by Dr. Turok and Dr. Oriti that allowed me to spend an intensive week at Perimeter Institute in October 2008 and several months at AEI Golm in 2009/10 deserve special thanks.

I am grateful to Trinity College for all its support, both financial and otherwise, over the course of the last four years, without which I would not have been able to complete a PhD at Cambridge. My work has also been supported by an EPSRC studentship.

I would like to thank my friends in Cambridge, who have hopefully improved my understanding of physics and made my time more enjoyable, not least through countless hours of ‘foosball’, and my friends in Germany who have kept in touch while I was away for so long, and gave support in difficult times.

Going back a bit further in time, I would not have ended up as a PhD student in Cambridge (and certainly not at Trinity) without the encouragement of Dipl.-Phys. Igor Pikovski and the strong support of Profs. Ingo Peschel and Jürgen Bosse. The latter two together with Hermann Schulz (Hochschuldozent) were also influential in shaping my interest in theoretical physics early on.

Last but not least, I would like to thank my aunt Odina and uncle Uwe for their support, especially for the first year at Cambridge.

Geometric Aspects of Gauge and Spacetime Symmetries

STEFFEN CHRISTIAN MARTIN GIELEN

Summary

We investigate several problems in relativity and particle physics where symmetries play a central role; in all cases geometric properties of Lie groups and their quotients are related to physical effects.

The first part is concerned with symmetries in gravity. We apply the theory of Lie group deformations to isometry groups of exact solutions in general relativity, relating the algebraic properties of these groups to physical properties of the spacetimes. We then make group deformation local, generalising deformed special relativity (DSR) by describing gravity as a gauge theory of the de Sitter group. We find that in our construction Minkowski space has a connection with torsion; physical effects of torsion seem to rule out the proposed framework as a viable theory. A third chapter discusses a formulation of gravity as a topological BF theory with added linear constraints that reduce the symmetries of the topological theory to those of general relativity. We discretise our constructions and compare to a similar construction by Plebanski which uses quadratic constraints.

In the second part we study CP violation in the electroweak sector of the standard model and certain extensions of it. We quantify fine-tuning in the observed magnitude of CP violation by determining a natural measure on the space of CKM matrices, a double quotient of $SU(3)$, introducing different possible choices and comparing their predictions for CP violation. While one generically faces a fine-tuning problem, in the standard model the problem is removed by a measure that incorporates the observed quark masses, which suggests a close relation between a mass hierarchy and suppression of CP violation. Going beyond the standard model by adding a left-right symmetry spoils the result, leaving us to conclude that such additional symmetries appear less natural.

Contents

1	Introduction	7
1.1	Lie Groups and Homogeneous Spaces	11
1.2	Structure of this Thesis and Conventions	15
I	Deformed and Constrained Symmetries in Gravity	17
2	Deformations of Spacetime Symmetries	18
2.1	Overview	18
2.2	Deformation Theory	21
2.3	The Petrov and Kaigorodov-Ozsváth Solutions	25
2.4	The Kaigorodov Solution and Lobatchevski Plane Gravitational Waves	42
2.5	Summary	49
3	Deformed General Relativity and Torsion	50
3.1	Introduction	50
3.2	Curved Momentum Space	55
3.3	Deformed Special Relativity	58
3.4	A de Sitter Gauge Theory of Gravity	60
3.5	Cartan Connections on Homogeneous Bundles	67
3.6	Synthesis	71
3.7	Summary and Discussion	74
3.8	Gauge Invariance Broken?	77
3.9	Symmetric Affine Theory	78
4	General Relativity From a Constrained Topological Theory	80
4.1	Overview	80
4.2	Actions for Gravity	82

4.3	Constraining BF Theory	85
4.4	Linear Constraints for BF-Plebanski Theory	90
4.5	Lagrangian and Hamiltonian Formulation	101
4.6	Summary and Outlook	103
II	Gauge Symmetries and CP Violation	106
5	$SU(3)$ and its Quotients	107
5.1	Introduction	107
5.2	The Homogeneous Space $SU(3)/U(1)^2$	111
5.3	Metrics on $U(1)^2 \backslash SU(3)/U(1)^2$	112
5.4	Hermitian and Complex Matrices	114
6	Naturalness of CP Violation	120
6.1	Introduction and Motivation	120
6.2	CP Violation in the Standard Model and Beyond	124
6.3	Statistics on the Space of CKM Matrices	129
6.4	Statistics on the Space of Mass Matrices	136
6.5	Extension to Neutrinos	151
6.6	Summary and Outlook	154
	Summary and Outlook	157
	References	161
	List of Figures	176
	List of Tables	177
	Index	178

Chapter 1

Introduction

“Ich kann es nun einmal nicht lassen, in diesem Drama von Mathematik und Physik – die sich im Dunkeln befruchten, aber von Angesicht zu Angesicht so gerne einander verkennen und verleugnen – die Rolle des (wie ich genugsam erfuhr, oft unerwünschten) Boten zu spielen.” [164]

Hermann Weyl

It is impossible to overemphasise the importance of symmetry as a fundamental concept in both mathematics and theoretical physics.

In mathematics, the study of groups arose as an abstraction of “transformation groups”, invertible maps of a given “point field” into itself [164]. Groups are the mathematical realisation of the concept of symmetry. Felix Klein’s Erlangen Programme [96] aimed to describe any geometry by its transformation group; as we will detail shortly, the geometries considered in the Erlangen Programme are given by coset spaces of the form $G/H = \{gH, g \in G\}$, where G is a continuous (Lie) group and H a closed subgroup of G . Klein himself found the concept of groups most characteristic of 19th century mathematics [164]. The study of continuous symmetries is, in mathematical terms, the study of Lie groups and homogeneous spaces which nicely combines aspects of group theory and differential geometry. Beyond the direct study of groups, symmetry is an important concept in other parts of mathematics, such as the study of differential equations.

In physics, the underlying symmetries of any theory are arguably its most fundamental property. As an example, the first physical theory in the modern sense, Newtonian mechanics, fundamentally rests on assumptions about the symmetries of

the physical world: A configuration of bodies in (gravitational) interaction will behave in the same way if it is rotated as a whole (space is isotropic); a common motion of all bodies at constant velocity is not observable (relativity of inertial frames). Consequently, the laws of Newtonian mechanics are invariant under the action of the semidirect product of the rotation group $SO(3)$ and commutative Galileian boosts, the **Euclidean group** $E(3)$. When Einstein understood in 1905, in the wake of the Michelson-Morley experiment [118], that it was necessary to modify the concept of relativity of inertial frames, he did so by replacing the Euclidean group by the Lorentz group $SO(3, 1)$ as the fundamental symmetry group. We will see in chapter 2 that this process can be mathematically viewed as a Lie group (or algebra) deformation.

The Galilean and Lorentz groups are examples of **spacetime** symmetries; their primary action on space(time) itself induces an action on all fields according to their tensorial structure. In general relativity, these global spacetime symmetries are made local, although global symmetries also appear in many exact solutions of physical interest, as we shall see in chapter 2. Both global and local spacetime symmetries will play a role in the first part of this thesis.

More generally, the equations defining a given theory (or the action, in the usual case that the equations of motion can be derived from an action principle) may also be invariant under a group of transformations of the physical variables which is not associated to a transformation of spacetime. A simple example in field theory is that of a complex scalar field ϕ with equations of motion invariant under a global rotation $\phi \rightarrow e^{i\alpha}\phi$. Promoting such a global symmetry to a local symmetry leads to a **gauge** theory. We will discuss gauge symmetries in both gravity and particle physics in this thesis.

Mathematically, the fundamental structures describing local symmetries are principal fibre bundles over a given manifold which is interpreted as spacetime, where the structure group becomes the “gauge group” of the theory, and a connection on such a bundle is interpreted as a gauge potential. The remaining fields that make up the physical content of the theory then live in associated vector and tensor bundles. While this underlying structure suggests a very close relationship of all fundamental interactions in Nature, there are important differences between Yang-Mills theory, defined in terms of a principal bundle with compact gauge group, and general relativity, where in addition to the connection one has a frame field as a second structure

which is crucial for the geometric interpretation of the theory.* In chapter 3, we will study a formulation of general relativity as a Yang-Mills-like theory which tries to exploit this analogy; the usual Lorentz algebra valued connection and the frame field are combined into a connection on a bundle of the Poincaré group (or its deformation, the (anti-)de Sitter group). This study will elucidate some of the virtues as well as shortcomings of such a formulation.

Symmetry is our most important guiding principle towards discovering the laws of Nature. In particle physics, the electroweak model [72, 146, 161] unifying the electromagnetic and weak interactions was discovered by postulating the fundamental symmetries and writing down the most general action invariant under them. A crucial step in completing the present picture of particle physics was the proof by 't Hooft and Veltman [156] that the mathematical structure of Yang-Mills theory is compatible with the framework of quantum field theory; namely, Yang-Mills theory is renormalisable. By far most attempts to go beyond the standard model of particle physics, such as supersymmetric theories and GUTs, are based on postulating new symmetries. Part of our analysis of chapter 6 will be to discuss possible observable consequences of such an existence of additional symmetries beyond those of the electroweak model.

General relativity, however, has so far evaded a similar treatment in terms of conventional field theory formulated on a background Minkowski spacetime, and by now it seems unlikely that adding (super)symmetries will be sufficient for a consistent quantum-mechanical treatment, although there has recently been increasing interest in the possibility that maximal ($N = 8$) supergravity in four dimensions is perturbatively finite [22]. String theory is one attempt to go beyond standard quantum field theory by replacing the notion of point particles by extended objects; symmetries (such as dualities) play a fundamental theory in string theory as well. Besides the symmetries that will be important for a consistent fundamental theory of quantum gravity, if we manage to construct (and understand) one, one may wonder whether at low energy some effects of quantum gravity will be manifest in a deformation of spacetime symmetries. This idea, which has been proposed on a purely phenomenological basis [4] as well as a consequence of developments in non-perturbative quantum gravity [6], will be taken up in chapter 3.

*A second difference is of course the non-compactness of the gauge group for Lorentzian signature, but this seems a less important difference to Yang-Mills theory in the basic structure of the theory – one often considers Riemannian signature as well.

The notion of symmetries we have discussed so far refers to invariance of the action or equations of motion under certain group actions. From a canonical viewpoint, there is another important class of symmetries, namely those generated by constraints. To obtain the canonical formulation of a given theory, one performs a Legendre transformation, leading to a Hamiltonian

$$p_i(x) \equiv \frac{\delta S}{\delta \dot{q}^i(x)}, \quad \mathcal{H} \equiv \int d^{d-1}x \mathcal{H}(x) \equiv \int d^{d-1}x (p_i(x) \dot{q}^i(x) - \mathcal{L}(x)), \quad (1.0.1)$$

where $S = \int dt \int d^{d-1}x \mathcal{L}(x)$ is the action defining the theory, and a symplectic structure on the phase space parametrised by (q^i, p_i) ,

$$\{q^i(x), p_j(y)\} = \delta_j^i \delta^{d-1}(x - y), \quad \{q^i(x), q^j(y)\} = \{p_i(x), p_j(y)\} = 0. \quad (1.0.2)$$

The Hamiltonian density $\mathcal{H}(x)$ typically splits into a “dynamical” part which is a generic function of the p_i and q^i , and a “constraint” part of the form $\lambda^\alpha f_\alpha(p_i, q^i)$; here λ^α are Lagrange multipliers, variables that appear in the original action whose canonical momentum is identically zero and that are conventionally not treated as belonging to the phase space. The functions $f_\alpha(p_i, q^i)$ then act as **constraints** imposed on the dynamical variables. A subset of these will be *first class*[†], *i.e.* their Poisson brackets vanish up to a linear combination of constraints. The Poisson brackets $\{f_\alpha, \cdot\}$ of the first class constraints then determine vector fields on the phase space whose integral curves are interpreted as describing physically equivalent configurations. Although not necessarily associated with connections on fibre bundles, the transformations on the phase space generated by these vector fields are, by their physical interpretation, usually interpreted as “gauge transformations” as well. We will see an instance of such a symmetry in the context of a topological field theory, which is then constrained by a term that is added to the action, in chapter 4.

Although our discussions of chapter 4 will remain classical, the motivation for the reformulation of general relativity that we will consider comes from the fact that the quantisation of topological field theories is relatively well understood. One might therefore hope that starting from a topological quantum theory which is then constrained may provide a different route to quantum general relativity. This also provides a bridge to the discussion of chapter 3 in that both consider possible reformulations of general relativity at the classical level which are suggested by developments in non-perturbative approaches to quantising gravity.

[†]We are here and in the following using the terminology introduced by Dirac [46].

In all of our discussions, an important distinction is that between discrete and continuous symmetries. The first systematic study of groups in physics was probably that of discrete groups describing the symmetries of crystals [164]. We shall be mainly interested in continuous symmetries which are described by Lie groups. The spacetime and gauge symmetries described above are of this type. Their differentiable structure means that differential geometry can be used to study them. In particular, we will see that the group structure is encoded in a bracket operation on the tangent space to the group at the identity element.

We will also encounter discrete symmetries, such as charge conjugation and parity in particle physics which are represented by the finite group \mathbb{Z}_2 . These will not be studied mathematically, but their physical significance will be discussed.

In the spirit of Hermann Weyl's self-image as a "messenger" between mathematics and physics, we will highlight different examples where a study of the differential geometry of (spacetime or gauge) symmetry groups leads to important physical consequences. Let us first introduce the most important aspects of the theory of Lie groups and homogeneous spaces that will be important throughout this thesis.

1.1 Lie Groups and Homogeneous Spaces

The study of symmetries is the study of groups. If we consider continuous symmetries, the corresponding groups have an additional differentiable structure: They are manifolds as well as groups[‡]. As simple examples one may think of the groups of isometries of \mathbb{R}^n , \mathbb{C}^n or \mathbb{H}^n , denoted by $O(n)$, $U(n)$ and $Sp(n)$ respectively, or more generally, of groups of matrices with continuous parameters which are submanifolds of $\mathbb{R}^{n \times n}$, $\mathbb{C}^{n \times n}$ or $\mathbb{H}^{n \times n}$. All examples of relevance in this thesis will be of this type.

Such groups are known as **Lie groups**, named after the Norwegian mathematician Sophus Lie (1842-1899) who initiated their systematic study in the context of systems of differential equations. Lie himself seems to have referred to Lie groups as "continuous groups" which describes appropriately what they are. The study of Lie groups combines aspects of differential geometry with group theory.

Two natural actions of a Lie group G on itself are given by the left and right

[‡]One also requires the group multiplication map $(x, y) \mapsto x \cdot y$ and the inversion map $x \mapsto x^{-1}$ to be differentiable.

multiplication maps

$$L : G \times G \rightarrow G, \quad (h, g) \mapsto hg, \quad R : G \times G \rightarrow G, \quad (h, g) \mapsto gh. \quad (1.1.1)$$

By definition of a group these maps act simply transitively: Given $a, g \in G$ there is a unique element h such that $L(h, g) = a$ and a unique element h' such that $R(h', g) = a$. The pull-back and push-forward of these maps act on the tangent and cotangent bundle of G .

Starting with a basis of the tangent space at the identity element e , one can use the push-forward of left translation to define a basis of vector fields on G , known as **left-invariant vector fields**. These define a global frame field on G and hence every Lie group is parallelisable. (Right-invariant vector fields can be defined analogously but often left invariance is preferred.) The tangent space at e , or alternatively the set of left-invariant vector fields on G , is known as the **Lie algebra** \mathfrak{g} . It inherits a **Lie bracket** $[\cdot, \cdot]$ from the commutator of vector fields. For matrix Lie groups, the bracket is given by the commutator of matrices.

Lie's important results are summarised in the following theorem [9]:

Theorem 1.1.1 (*Lie*)

- (1) For any Lie algebra \mathfrak{g} there is a Lie group G (not necessarily unique) whose Lie algebra is \mathfrak{g} .
- (2) Let G be a Lie group with Lie algebra \mathfrak{g} . If H is a Lie subgroup of G with Lie algebra \mathfrak{h} , then \mathfrak{h} is a subalgebra of \mathfrak{g} . Conversely, for each Lie subalgebra \mathfrak{h} of \mathfrak{g} , there exists a unique connected Lie subgroup H of G which has \mathfrak{h} as its Lie algebra. Furthermore, normal subgroups of G correspond to ideals in \mathfrak{g} .
- (3) Let G_1, G_2 be Lie groups with corresponding Lie algebras $\mathfrak{g}_1, \mathfrak{g}_2$. Then if \mathfrak{g}_1 and \mathfrak{g}_2 are isomorphic as Lie algebras, then G_1 and G_2 are locally isomorphic. If the Lie groups G_1, G_2 are simply connected, then G_1 is isomorphic to G_2 .

Loosely speaking, all of the structure of the group G is encoded in its Lie algebra. If one is not interested in the topological properties of G and considers only connected and simply connected Lie groups, one can study the geometry of G by studying \mathfrak{g} which is just a linear vector space together with a bracket structure.

For matrix Lie groups, it is usually most convenient to fix a basis of generators M_a and to expand the **Maurer-Cartan form** in terms of this basis:

$$g^{-1} dg = \sigma^a M_a \quad (1.1.2)$$

determines a basis of left-invariant one-forms σ^a . Similarly, right-invariant forms can be obtained by computing $dg g^{-1}$ which is clearly invariant under $g \mapsto gh$. The basis dual to $\{\sigma^a\}$ is then a basis of left-invariant vector fields. By taking the exterior derivative of both sides of (1.1.2) and using $d(g^{-1}) = -g^{-1} dg g^{-1}$ for matrices, one can show that the forms σ^a satisfy the Maurer-Cartan relations

$$d\sigma^a = -\frac{1}{2}C_b^a{}_c \sigma^b \wedge \sigma^c, \quad (1.1.3)$$

where $C_a^c{}_b$ are the **structure constants** of the Lie algebra defined by

$$[M_a, M_b] = C_a^c{}_b M_c. \quad (1.1.4)$$

One defines left-invariant tensors accordingly, *e.g.* left-invariant metrics as generated by symmetrised combinations of σ^a : Any metric of the form

$$g = h_{ab} \sigma^a \otimes \sigma^b \quad (1.1.5)$$

for constant symmetric h_{ab} is by definition left-invariant. One can phrase this by saying that there is a one-to-one correspondence between scalar products on the Lie algebra \mathfrak{g} (specified by h_{ab}) and left-invariant metrics on G [9]. These metrics are the ones compatible with the left action on G on itself.

Lie groups are in some sense special cases of a more general class of manifolds that will be important in the following: In many situations of physical interest, one considers a manifold X which a Lie group G acts on transitively, but not necessarily simply transitively. That is, for any $x, x' \in X$, there is a (not necessarily unique) $g \in G$ with $g \cdot x = x'$. (Again we take this to be a left action, the case of a right action is analogous.) Then X is called a **homogeneous space**. For a given $x \in X$, the subgroup $G_x = \{g \in G : g \cdot x = x\}$ leaving x invariant is called the *stabiliser*, *little group* or *isotropy subgroup* of x . Then we have the following result [9]

Proposition 1.1.2 *Let $G \times X \rightarrow X$ be a transitive action of a Lie group G on a manifold X , and let $H = G_x$ be the isotropy subgroup of a point x . Then:*

- (a) *The subgroup H is a closed subgroup of G .*
- (b) *The natural map $j : G/H \rightarrow X$ given by $j(gH) = g \cdot x$ is a diffeomorphism. In other words, the orbit $G \cdot x$ is diffeomorphic to G/H .*
- (c) *The dimension of G/H is $\dim G - \dim H$.*

Since we assume transitivity of the group action, the orbit $G \cdot x$ is the whole of X ; one can then identify X with the coset space G/H .[§] The choice of x in this identification is arbitrary and corresponds to a choice of origin in X ; different choices for x lead to groups G_x related by conjugation. The natural left action of G on G/H is just induced by the group multiplication law.

In many applications, an additional assumption is satisfied which simplifies the construction of metrics on a homogeneous space: A given homogeneous space G/H is called **reductive** if there is a splitting of Lie algebras of the form

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{x} \tag{1.1.6}$$

invariant under the adjoint action of H . Then the subspace \mathfrak{x} of the Lie algebra \mathfrak{g} is isomorphic [9] to the tangent space to G/H at H (or any other coset gH). It therefore defines a basis of left-invariant vector fields on X , and one can take the dual basis of a basis for \mathfrak{x} to construct left-invariant metrics on X . Reductive geometries will be of importance in chapter 3 (with $G = SO(d, 1)$ and $H = SO(d - 1, 1)$) and in chapters 5 and 6 (with $G = SU(3)$ and $H = U(1)^2$). We will encounter an example of a non-reductive geometry in section 2.4.

Homogeneous spaces are the simplest generalisations of (flat) Euclidean geometry: In the 19th century, attempts to deduce Euclid's fifth postulate, the parallel postulate [50], from the other four[¶], led to the discovery of elliptic and hyperbolic geometries based on the sphere $S^n \simeq SO(n + 1)/SO(n)$ and hyperbolic space $H^n \simeq SO(n, 1)/SO(n)$, the natural counterparts of Euclidean space $\mathbb{R}^n \simeq E(n)/SO(n)$. In Lorentzian geometry, the spaces of constant curvature, Minkowski, de Sitter and anti-de Sitter space, are also homogeneous spaces. Many other known exact solutions in general relativity are also homogeneous spaces; see [154, chapter 12].

Homogeneous spaces appear whenever there are (gauge or spacetime) symmetry groups. Another example would be a Higgs potential spontaneously breaking a symmetry group G down to a subgroup H ; then the space of vacua would be G/H .

[§]This result justifies dropping the distinction between the terms “coset space” and “homogeneous space” as we will do in the following.

[¶]About 30 such “proofs” were shown to be unsatisfactory in the 1763 dissertation of A. Kaestner's student Georg Simon Klügel entitled “Conatuum praecipuorum theoriam parallelarum demonstrandi recensio” (Critique of the foremost attempts to prove the theory of parallels) [97, p.274].

1.2 Structure of this Thesis and Conventions

The thesis consists of two main parts.

The first part explores symmetries in relativity which are deformed or constrained. In chapter 2 we introduce deformation theory of Lie algebras and investigate its application to global symmetries (isometries) of certain exact solutions to Einstein’s equations in four dimensions. We will see how the group-theoretic viewpoint of spacetimes in terms of their isometry groups relates to physical properties such as causality violation. Section 2.3 contains calculations and results taken from [65] and are the result of a collaboration with Gary Gibbons, except subsection 2.3.6 which is new. We omit those parts of [65] which are of less relevance to the discussion of symmetry groups. In chapter 3 we turn to a study of local symmetries in relativity, namely those of the tangent space. We investigate how Cartan geometry, a generalisation of Riemannian geometry where the flat tangent space is replaced by an arbitrary homogeneous space, can be used to generalise the proposed deformed special relativity (DSR) to a “deformed general relativity” in the same way that (pseudo-)Riemannian geometry describes general relativity as an extension of special relativity. We analyse the physical predictions of a particular example supposedly describing Minkowski space, and note that it involves a connection with torsion potentially leading to disastrous predictions of charge nonconservation. This analysis serves as another example of the interplay between an algebraic viewpoint on (infinitesimal) symmetries and physical effects. The material presented in chapter 3 is mainly taken from [66] which again resulted from collaboration with Gary Gibbons. In chapter 4 we study yet another possible reformulation of classical general relativity as a topological BF theory with constraints. Such a formulation has been considered before and is of relevance in covariant approaches to quantising gravity, but we give a new version of it where the constraints are linear, removing some of the ambiguity in their solutions. We see how a very large symmetry on the phase space of the theory – the local equivalence of all solutions to the equations of motion – can be constrained to give general relativity, a theory with local degrees of freedom. The results detailed in chapter 4, published in [69], were obtained in collaboration with Daniele Oriti.

The second part is concerned with gauge symmetries in particle physics and their influence on CP violation in the electroweak sector. We aim to give estimates for the likelihood of the observed magnitude of CP violation, measured by the Jarlskog invariant J , by constructing a natural measure on the space of Cabibbo-Kobayashi-

Maskawa (CKM) matrices. This space is a double quotient of $SU(3)$ and does not have a clearly preferred measure. After introducing different choices for the measure in section 5, motivated first only by geometric and then also by physical considerations – we will incorporate the observed quark masses into the analysis as well – in chapter 6 we make a detailed study of their physical predictions. We find that once the quark masses are assumed, predictions for CP violation agree well with observation. We then provide another example of how modified symmetries affect physical predictions by assuming an extended left-right symmetry given by a second $SU(2)$ gauge group. This will influence our predictions on CP violation and lead to completely different results. The analysis for the standard model was done in collaboration with Gary Gibbons, Chris Pope and Neil Turok in [67]. Much of the calculations presented in sections 6.3 and 6.4 are taken from this paper; I have omitted some parts of [67] where my original contribution has been limited. The calculations in subsection 6.4.2 are my own and were published in [68].

We close with a brief summary and outline possible directions for future research.

Our conventions for general relativity are those of Hawking & Ellis; the signature of a Lorentzian metric is mainly positive; Greek letters are used for spacetime and Latin letters for internal indices. The curvature of a connection is always $F = dA + A \wedge A$ without any coupling constants as customary in some of the particle physics literature. In discussions of CP violation in particle physics we follow the book of Jarlskog [91].

Part I

Deformed and Constrained Symmetries in Gravity

Chapter 2

Deformations of Spacetime Symmetries

2.1 Overview

The archetypal examples of groups describing spacetime symmetries in physics are the Galilean and Lorentz groups of classical mechanics and special relativity. The crucial insight of Einstein in 1905 was the necessity of replacing the Galilean group by the Lorentz group in order to have an invariant speed of light c ; as we will see shortly, this process is also the archetypal example of a **Lie algebra deformation**. Historically, the development of a mathematical theory of deformation and contraction of Lie algebras was motivated by advances in theoretical physics, but it is at least conceivable that the replacement of the Galilean group by the Lorentz group had been suggested by mathematicians working on Lie group theory, instead of physicists like Einstein.

Since physical theories are to some extent always idealised approximations of a much more complex Nature, the process of deformation of a symmetry group allows one to make statements about the stability of a given theory. This stability point of view was explored in detail in [116]. The passage from non-relativistic to special-relativistic mechanics can be understood as a deformation of an “unstable” to a “robust stable” theory [116]. The point here is that the zero commutators of Galilean boosts constitute “fine-tuning” and any perturbation of the commutation relations leads to noncommuting boosts. On the other hand, given noncommuting boosts, a slight perturbation of the coefficient appearing in the commutator can be undone by redefining the algebra generators, hence one has “robustness” in the theory.

In [116], the possible central extension of the Poisson algebra of the phase-space coordinates of Hamiltonian classical mechanics, leading to quantum mechanics, was interpreted as a similar instability of the original algebra. It was then argued that although the resulting Heisenberg algebra admits a deformation, it can be undone by a nonlinear redefinition of the position coordinate.

In this chapter, we shall focus on Lie algebra deformations, perturbations of the structure constants which can be expanded in a power series. We are maintaining the structure of a Lie algebra, *i.e.* of a vector space, and redefinitions of the generators will always be linear transformations. One could consider more general kinds of deformations, where the Lie algebra structure is lost, either by allowing commutators of two generators to be nonlinear in the generators or by allowing “structure functions”.

While the existence or non-existence of deformations of a given Lie algebra is a mathematically well-posed question, the physical interpretation of its answer is less clear in general. Both the generators and the deformation parameters of a given Lie algebra may have different possible interpretations. For example, while the Poincaré group admits a deformation to the de Sitter or anti de Sitter groups, it was argued in [116] that the deformation parameter R , defining a length scale, may be taken to be infinity if one is only interested in local kinematics, which of course is what happens in general relativity. It is also clear that while linear transformations can formally be applied to the generators in the Lie algebra, it may not appear physically meaningful to take linear combinations of, for instance, translation and rotation generators in the Poincaré algebra. This point was stressed in [37], who also pointed out that the interpretation of position operators appearing in a “kinematical algebra” acting on single- or multi-particle quantum states is physically rather unclear, since position is not additive in a way that momentum and angular momentum are.

The classic cases of the de Sitter group and its contractions were studied in detail in [12], as we will review shortly. One could argue that the mathematical theory behind deformations seems to have predictive power in terms of extensions or generalisations of fundamental theories, but these are only *a posteriori* predictions [37]. There seem to be two obvious classes of new applications for the concept of deformations of spacetime symmetries:

- Situations where the relevant symmetry group is a subgroup of the Poincaré group, or more generally a group of lower dimension than the Poincaré group. One example for the former is the proposed “very special relativity” [39], where

the fundamental symmetry group is taken to be $\text{ISIM}(n-2)$, consisting of the maximal subgroup of the Lorentz group leaving a null direction invariant plus translations. The possibility of deforming this group to obtain “general very special relativity” was discussed in [62]; a physically acceptable one-parameter deformation was found, similar to the case of the Poincaré group which can be deformed into the (anti-)de Sitter group. Interestingly enough, this deformed group turns out to be the symmetry group of an asymptotically AdS spacetime playing a role in the construction of nonrelativistic hydrodynamics in the context of AdS/CFT: The “Schrödinger spacetime” [81]

$$ds^2 = r^2 \left(-2 du dv - r^{2\nu} du^2 + d\vec{x}^2 \right) + \frac{dr^2}{r^2} \quad (2.1.1)$$

can be viewed as a left-invariant metric on a group manifold which is a subgroup of the deformed group $\text{DISIM}_b(n-2)$ [63, app. B].

On a less fundamental level, given a homogeneous space G/H with isometry group G , *e.g.* as a given solution of general relativity, one may always ask what deformations its isometry group admits, and try to construct spaces with the deformed symmetry. The two examples we give in this chapter describe this kind of situation. In principle, this could be used to facilitate the search for new solutions of general relativity, although one is restricted to homogeneous spaces, and curvature invariants are constant. One may still obtain spaces with interesting causal properties, as exemplified by the BTZ black hole [152] (obtained from anti-de Sitter space by periodic identifications), and by the closed timelike curves that we will see occur in the Petrov and Kaigorodov-Ozsváth solutions.

- Alternatively, one may assume that the symmetry group is larger than the Poincaré group. In particular, the formalism can be straightforwardly extended to supergroups [52]. It was found in [52] that the Killing superalgebra of the M2-brane in eleven dimensions admits a deformation, which leads to the conjecture that a solution with the deformed Killing superalgebra exists, possibly having an interpretation as a perturbation of the M2-brane. Unfortunately, the search for such a solution does not seem to have been successful*.

Another example is Peter West’s conjecture of E_{11} as being the relevant symmetry group for M-theory [162], where one could look for possible deformations.

*José Figueroa-O’Farrill, private communication (May 2009)

As outlined in section 1.1, the general setup for our considerations is a manifold X which a group of transformations G acts on transitively, which may then be identified with G/H where H is a subgroup of G .

When using deformations to look for interesting manifolds “close to” a given one, one considers group deformations of the group G and then identifies the appropriate stabiliser H in the deformed group. It simplifies the interpretation of the manifold one obtains if this stabiliser can be viewed as a deformation of the stabiliser subgroup of the original undeformed group, but in general this need not be the case [62].

By definition, the deformation theory reviewed here can only make statements about Lie algebras. As explained in section 1.1, Lie’s theorems allow us to use the terms “Lie algebra” and “Lie group” interchangeably, meaning by “Lie group” the unique connected, simply connected group specified by a given algebra. In particular, terms such as “Poincaré group” or “de Sitter group” refer to connected components of the identity.

We review the mathematical theory of Lie algebra deformations in section 2.2, and then give two examples for pairs of solutions of general relativity related by such deformations: The Petrov and Kaigorodov-Ozsváth solutions are discussed in section 2.3. We shall see that they are both geodesically complete, and try to relate their causal properties. We then turn to the Kaigorodov solution in section 2.4, and find it can be deformed into a solution known as a “Lobatchevski plane gravitational wave”.

2.2 Deformation Theory

The theory of deformation of Lie algebras was reviewed in [104], where it was connected to the perhaps more familiar operation of Lie algebra (Inonu-Wigner [88]) contraction. A Lie algebra \mathfrak{g} is deformed by redefining the Lie brackets as a power series in a parameter t

$$f_t(a, b) = [a, b] + tF_1(a, b) + t^2F_2(a, b) + \dots, \quad a, b \in \mathfrak{g}, \quad (2.2.1)$$

where the series is required to converge in some neighbourhood of the origin. The Jacobi identity then leads to integrability conditions on the functions F_i at each order in t :

$$\sum_{\mathcal{P}(a,b,c)} \sum_{\mu+\nu=n} F_\mu(F_\nu(a, b), c) + F_\nu(F_\mu(a, b), c) \stackrel{!}{=} 0, \quad (2.2.2)$$

where the sum is over all cyclic permutations and $n = 0, 1, 2, \dots$

Choosing $n = 0$ gives back the Jacobi identity for $[\cdot, \cdot]$, and at linear order ($n = 1$) one gets

$$\sum_{\mathcal{P}(a,b,c)} F_1([a, b], c) + [F_1(a, b), c] \stackrel{!}{=} 0. \quad (2.2.3)$$

A deformation that only corresponds to a change of basis in the algebra \mathfrak{g} will be regarded as trivial; these are of the form

$$f(a, b) = S[S^{-1}a, S^{-1}b], \quad S \in GL(n), \quad (2.2.4)$$

and expanding S in a power series around the identity gives the general form of trivial deformations at each order in t .

It is perhaps more common to express these operations in terms of *structure constants* $C_a^c{}_b$, as defined in (1.1.4) for a given basis. Denoting the structure constants of the deformed algebra, with respect to the same basis, by $\hat{C}_a^c{}_b$, the Jacobi identity for the deformed algebra can be written as

$$\hat{C}_d^e{}_a \hat{C}_b^d{}_c = 0. \quad (2.2.5)$$

One can then contemplate $\mathbb{R}^{n^2(n-1)/2}$ (where n is the dimension of \mathfrak{g}) parametrised by coordinates $\hat{C}_a^c{}_b$ antisymmetric in the lower indices, and the submanifold of $\mathbb{R}^{n^2(n-1)/2}$ defined by the Jacobi identity. A deformation of the given algebra \mathfrak{g} describes a smooth curve in this submanifold which can be expanded in a power series in a parameter t :

$$\hat{C}_a^c{}_b(t) = C_a^c{}_b + tA_a^c{}_b + t^2B_a^c{}_b + \dots \quad (2.2.6)$$

At linear order in t the Jacobi identity (2.2.5) holds if

$$C_d^e{}_a A_b^d{}_c + A_d^e{}_a C_b^d{}_c = 0. \quad (2.2.7)$$

Then a linear deformation only gives rises to a deformation if the requirement (2.2.5) can be satisfied at each order in t .

Under a change of basis in \mathfrak{g} , expressed as a matrix S , the structure constants will change according to

$$\hat{C}_a^b{}_c(t) = S^b{}_e C_d^e{}_f (S^{-1})^d{}_a (S^{-1})^f{}_c. \quad (2.2.8)$$

Expanding $S^a{}_b(t) = \delta_b^a + tM^a{}_b + \dots$, this means that to first order a trivial deformation can be written as

$$A_a^b{}_c = M^b{}_e C_a^e{}_c - C_e^b{}_c M^e{}_a - C_a^b{}_e M^e{}_c. \quad (2.2.9)$$

One can rephrase these conditions in the language of differential forms by choosing a basis $\{\lambda^a\}$ of left-invariant 1-forms for the original algebra satisfying the Maurer-Cartan relations $d\lambda^a = -\frac{1}{2}C_b{}^a{}_c\lambda^b \wedge \lambda^c$, and also defining $A^a = \frac{1}{2}A_b{}^a{}_c\lambda^b \wedge \lambda^c$ to be a vector-valued 2-form and $C^a{}_b = C_c{}^a{}_b\lambda^c$ to be a matrix-valued 1-form [62]. Here $C_b{}^a{}_c$ are the structure constants of the original algebra and $A_b{}^a{}_c$ is their infinitesimal deformation. Then at linear order, the Jacobi identity is

$$DA := dA + C \wedge A = 0, \quad (2.2.10)$$

and a deformation that can be written as $A = DS$ for some $S = S^a{}_b\lambda^b$ is regarded as trivial. Hence non-trivial infinitesimal deformations are in one-to-one correspondence with the cohomology group $H^2(\mathfrak{g}, \mathfrak{g})$ [53]. At higher order one has to consider the higher cohomology groups. In general, if a non-trivial deformation has been found infinitesimally (*i.e.* at linear order), it does not necessarily extend to a deformation satisfying the full set of conditions (2.2.2); obstructions to “integrability” of infinitesimal deformations are given by the cohomology group $H^3(\mathfrak{g}, \mathfrak{g})$ [53]. In practice, once all infinitesimal deformations have been found, one can verify the full Jacobi identity, and if it is violated, try to add higher order terms to the deformation.

Using cohomology theory, one can determine all general deformations of a given algebra, as was done for the Galilean algebra, with and without central extension, in [53]. We will focus on deformations at linear order in the following.

Applying this formalism, the authors of [12] classified all possible kinematical groups in four dimensions, by which they meant classifying ten-dimensional Lie algebras consisting of “translations”, “rotations”, and “inertial transformations”. This splitting meant singling out three generators (“rotations”) and fixing their commutators with the other generators to assure isotropy of space. They further assumed that parity and time-reversal (defined on the generators) leave the Lie brackets invariant, and that the subgroup of inertial transformations is non-compact. The classification one then obtains consists of eight types of Lie algebras, all of which may be obtained by either contracting the de Sitter algebra $\mathfrak{so}(4, 1)$ or anti-de Sitter algebra $\mathfrak{so}(3, 2)$, or alternatively as deformations of the Abelian algebra, as shown in figure 2.1. The possible contractions correspond to different physical limits such as vanishing cosmological constant, infinite speed of light or zero speed of light, as explained in detail in [12]. For instance, the Newton groups describe the symmetries of a non-relativistic cosmological model with nonzero cosmological constant; this was discussed in detail in [64].

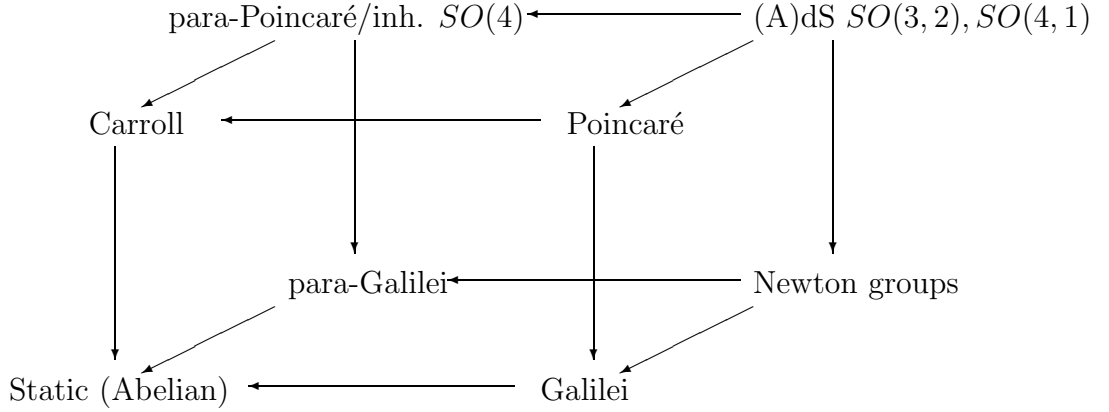


Figure 2.1: Possible contractions and deformations of kinematical groups, according to [12]. Arrows denote possible Inonu-Wigner [88] contractions.

Looking at figure 2.1, it could be argued that the introduction of noncommuting translations, the most direct effect of a curved spacetime, is very natural from the group deformation viewpoint as the de Sitter and anti-de Sitter groups arise as deformations of the Poincaré group.

For completeness, we mention that in order to obtain the algebra of quantum mechanics from a deformation of that of classical mechanics, as discussed in [116], one either deforms the Poisson algebra of functions on phase-space, replacing the Poisson bracket by a Moyal bracket

$$\{f, g\}_M = \{f, g\} - \frac{\hbar^2}{4 \cdot 3!} \sum_{\substack{i_1, i_2, i_3 \\ j_1, j_2, j_3}} \omega^{i_1 j_1} \omega^{i_2 j_2} \omega^{i_3 j_3} \partial_{i_1 i_2 i_3} (f) \partial_{j_1 j_2 j_3} (g) + \dots, \quad (2.2.11)$$

or alternatively considers the phase-space coordinates as elements of an Abelian Lie algebra which is replaced by the Heisenberg algebra. In the first case one has an infinite-dimensional algebra, in the second case one performs a central extension together with a deformation. The very illuminating discussion in [116] used purely mathematical arguments as well as physical intuition to establish the stability of the Heisenberg algebra, considered as a test for the “robustness” of quantum mechanics. We stress again that this is a different notion of deformation, since one is now considering non-linear transformations of the generators.

The maximal algebra one could consider in this discussion is the algebra of the Poincaré group, together with “spacetime coordinates” and an “identity” (a central element) which appear as operators in quantum mechanics. In four dimensions, this

is a 15-dimensional group which is naturally deformed into $\mathfrak{so}(6 - t, t)$, where the original central element ceases to commute with the other generators in general. This maximally extended case is most obviously fraught with interpretational difficulties, as clearly illustrated in [37]. Formally, the deformation process naturally leads to non-commuting spacetime coordinates [116]. This proposal goes back to Snyder [151], whose algebra is however not the same as the one proposed in [116]. The idea of “spacetime being non-commutative” is very interesting physically, however since it is separate from the deformation theory of spacetime symmetries reviewed in this chapter, it will be discussed in more detail in chapter 3.

2.3 The Petrov and Kaigorodov-Ozsváth Solutions

Among the wide range of known exact solutions to Einstein’s field equations (the most comprehensive source is [154]), there are some with interesting symmetry properties. In this section we concentrate on the particularly simple case of a simply transitive group of motions, such that the stabiliser is trivial and there is a one-to-one correspondence between spacetime points and elements of the group of motions. Therefore the spacetime is not only a manifold but also a Lie group. One example, and in fact the only example among vacuum solutions without a cosmological constant, is provided by the Petrov solution (2.3.1). An analogous example with negative cosmological constant is provided by the Kaigorodov-Ozsváth solution (2.3.10).

The Petrov solution has a physical interpretation as the exterior solution of an infinite rigidly rotating cylinder. To gain an understanding of the physical properties of the Kaigorodov-Ozsváth solution, we compute the stress-energy tensor of the boundary theory in the context of the AdS/CFT correspondence from an expansion near the conformal boundary. We then give a matrix representation of the isometry group of the Petrov solution and its algebra and use it to construct left-invariant one-forms on the group which could be used to compute left-invariant metrics.

We apply the theory of Lie algebra deformations to relate the Petrov solution to the Kaigorodov-Ozsváth solution, and discuss the physical interpretation of this result. It is well known that the Petrov solution contains closed timelike curves (CTCs), so an obvious question is whether the Kaigorodov-Ozsváth solution exhibits similarly exotic properties. Since in the context of possible causality violation by CTCs a central question is whether these can be created by some process in a spacetime which

did not exhibit CTCs initially, we give a brief discussion of the possible appearance of CTCs by spinning up a rotating cylinder. This elementary example supports Hawking's chronology protection conjecture which asserts that the appearance of CTCs is forbidden by the laws of physics [77].

In a more detailed analysis of the physical properties of the Petrov and Kaigorodov-Ozsváth solutions, we show that they are geodesically complete. We discuss the global causality properties of the two spaces; while the Petrov solution is totally vicious, due to a theorem of Carter, a similar analysis for the Kaigorodov-Ozsváth solution remains inconclusive. The computation of the stress-energy tensor in AdS/CFT however suggests causal pathologies in the Kaigorodov-Ozsváth solution as well.

2.3.1 Vacuum Solutions With Simply-Transitive Groups of Motions

$\Lambda = 0$ - The Petrov Solution

The Petrov solution is introduced in [154] in the following theorem: *The only vacuum solution of Einstein's equations admitting a simply-transitive four-dimensional maximal group of motions is given by*

$$k^2 ds^2 = dr^2 + e^{-2r} dz^2 + e^r (\cos \sqrt{3}r (d\phi^2 - dt^2) - 2 \sin \sqrt{3}r d\phi dt), \quad (2.3.1)$$

where k is an arbitrary constant, which shall be set equal to one, and we have relabelled the coordinates compared to [154]. The solution was first given in [134] and also discussed in [43]. It describes a hyperbolic plane H^2 (the (r, z) -plane) with a timelike two-plane attached to each point.

The isometry group is generated by the Killing vector fields

$$T \equiv \frac{\partial}{\partial t}, \quad \Phi \equiv \frac{\partial}{\partial \phi}, \quad Z \equiv \frac{\partial}{\partial z}, \quad R \equiv \frac{\partial}{\partial r} + z \frac{\partial}{\partial z} + \frac{1}{2}(\sqrt{3}t - \phi) \frac{\partial}{\partial \phi} - \frac{1}{2}(t + \sqrt{3}\phi) \frac{\partial}{\partial t}, \quad (2.3.2)$$

which satisfy the algebra

$$[R, T] = \frac{1}{2}T - \frac{\sqrt{3}}{2}\Phi, \quad [R, \Phi] = \frac{1}{2}\Phi + \frac{\sqrt{3}}{2}T, \quad [R, Z] = -Z. \quad (2.3.3)$$

The isometry group contains three-dimensional subgroups of Bianchi types I and VII_h acting on timelike hypersurfaces, and the solution (2.3.1) is Petrov type I [154]. The first three Killing vectors obviously generate translations while the action of the one-parameter subgroup generated by R on spacetime is given by the integral curves of

R , solutions of

$$\frac{dx^\mu(\lambda)}{d\lambda} = R^\mu(x(\lambda)). \quad (2.3.4)$$

These integral curves have the form

$$\begin{aligned} x^\mu(\lambda) = & \left(r_0 + \lambda, z_0 e^\lambda, \phi_0 e^{-\frac{\lambda}{2}} \cos \frac{\sqrt{3}}{2} \lambda + t_0 e^{-\frac{\lambda}{2}} \sin \frac{\sqrt{3}}{2} \lambda, \right. \\ & \left. -\phi_0 e^{-\frac{\lambda}{2}} \sin \frac{\sqrt{3}}{2} \lambda + t_0 e^{-\frac{\lambda}{2}} \cos \frac{\sqrt{3}}{2} \lambda \right), \end{aligned} \quad (2.3.5)$$

where we label the coordinates by $x^\mu = (r, z, \phi, t)$.

The metric components $g_{\phi\phi}$ and g_{tt} become zero at certain values[†] of r , but as the determinant of the metric in (2.3.1) is always -1 , it is possible to extend the coordinates to infinite ranges and the coordinates (r, z, ϕ, t) define a global chart. The manifold is also time-orientable, as the vector field $t^\mu = (0, 0, \sin \frac{\sqrt{3}}{2} r, \cos \frac{\sqrt{3}}{2} r)$ defines a global arrow of time, though this amounts to $\frac{\partial}{\partial t}$ being future- as well as past-directed at some points (as well as spacelike at others).

Bonnor [25] pointed out that the solution can be viewed as a special case of the exterior part of a Lanczos-van Stockum solution [103, 155] describing an infinite cylinder of rigidly rotating dust. Since this allows a physical interpretation of (2.3.1), let us give the general solution for an infinite rigidly rotating dust cylinder, which in Weyl-Papapetrou form is given by [157]

$$ds^2 = H(\rho)(d\rho^2 + dz^2) + L(\rho)d\chi^2 + 2M(\rho)d\chi d\tau - F(\rho)d\tau^2, \quad (2.3.6)$$

where H , L , M , F are functions of the radial variable ρ containing two parameters a and R , interpreted as the angular velocity and radius of the cylinder respectively.

The high energy case $aR > \frac{1}{2}$, which contains closed timelike curves (CTCs), is of interest here, and with the choices $R = \sqrt{e}$ and $aR = 1$ the exterior solution is given by

$$\begin{aligned} H(\rho) &= \frac{1}{\rho^2}, \quad L(\rho) = -2\sqrt{\frac{e}{3}}\rho \sin \left(\sqrt{3} \log \frac{\rho}{\sqrt{e}} \right), \\ M(\rho) &= \frac{2}{\sqrt{3}}\rho \sin \left(\frac{\pi}{3} + \sqrt{3} \log \frac{\rho}{\sqrt{e}} \right), \quad F(\rho) = \frac{2}{\sqrt{3}e}\rho \sin \left(\frac{\pi}{3} - \sqrt{3} \log \frac{\rho}{\sqrt{e}} \right), \end{aligned} \quad (2.3.7)$$

where ρ is a radial coordinate in the exterior of the cylinder and so is restricted to

$$\rho \geq \sqrt{e}, \quad (2.3.8)$$

[†]these values of r can be shifted by a coordinate transformation $\phi \rightarrow \alpha\phi + \beta t$, $t \rightarrow -\beta\phi + \alpha t$ and hence have no coordinate-independent significance

and χ is an angular coordinate and periodically identified with period 2π ; τ and \tilde{z} are unconstrained. Applying the coordinate transformations

$$\tilde{z} = \sqrt{e} z, \rho = \sqrt{e} e^r, \chi = \frac{1}{\sqrt[4]{3}\sqrt{2e}} \left(\sqrt{2 + \sqrt{3}z} + \sqrt{2 - \sqrt{3}t} \right), \tau = \frac{1}{\sqrt[4]{3}}(z - t) \quad (2.3.9)$$

to the line element (2.3.6) indeed gives back (2.3.1). Hence if we adopt the interpretation of (2.3.1) as describing the exterior of a spinning cylinder, we restrict the coordinates to $r \geq 0$ and identify $\sqrt{2 + \sqrt{3}z} + \sqrt{2 - \sqrt{3}t}$ with $\sqrt{2 + \sqrt{3}z} + \sqrt{2 - \sqrt{3}t} + 2\pi\sqrt[4]{3}\sqrt{2e}$.

The general Lanczos-van Stockum solution has three linearly independent Killing vectors $\frac{\partial}{\partial\chi}$, $\frac{\partial}{\partial\tau}$ and $\frac{\partial}{\partial z}$ but a fourth Killing vector is only present in the special case $aR = 1$ because the algebraic invariants of the Riemann tensor are independent of ρ just in this case [26].

$\Lambda < 0$ - The Kaigorodov-Ozsváth Solution

A solution of the vacuum Einstein equations with *negative* cosmological constant which has a simply-transitive four-dimensional group of motions was first given by Kaigorodov [93] and rediscovered by Ozsváth [125]. It has the line element

$$ds^2 = -\frac{3}{\Lambda} dr^2 + e^{-2r}(dz^2 + 2 dt d\phi) + e^{4r} d\phi^2 - 2\sqrt{2}e^r dz d\phi. \quad (2.3.10)$$

This solution was also given in [154]. It is Petrov type III and the metric asymptotically (as $r \rightarrow -\infty$) approaches that of anti-de Sitter space. Obvious Killing vectors are

$$Z \equiv \frac{\partial}{\partial z}, \Phi \equiv \frac{\partial}{\partial \phi}, T \equiv \frac{\partial}{\partial t}, \quad (2.3.11)$$

and the metric (2.3.10) has a further isometry

$$r \rightarrow r + \lambda, z \rightarrow e^\lambda z, \phi \rightarrow e^{-2\lambda} \phi, t \rightarrow e^{4\lambda} t \quad (2.3.12)$$

which is generated by the fourth Killing vector $R \equiv \frac{\partial}{\partial r} + z\frac{\partial}{\partial z} - 2\phi\frac{\partial}{\partial \phi} + 4t\frac{\partial}{\partial t}$. The Killing vector fields satisfy the algebra

$$[R, Z] = -Z, [R, \Phi] = 2\Phi, [R, T] = -4T. \quad (2.3.13)$$

No analogous solution for a positive cosmological constant exists [125, 154]. Because of the similarity to (2.3.3) one would expect (2.3.13) to arise as a deformation of (2.3.3). Physically both algebras describe the isometries of vacuum solutions of Einstein's

equations, one with $\Lambda < 0$ and one with $\Lambda = 0$, and hence one might expect one of them to arise as some limit of the other.

This spacetime is also time-orientable, as the vector field $t^\mu = (0, -\frac{1}{\sqrt{2}}e^{3r}, -1, 1)$ defines a global arrow of time.

We can express (2.3.10) in coordinates corresponding to Poincaré coordinates on AdS (with $\Lambda = -3$)

$$ds^2 = \frac{d\rho^2 + dz^2 + d\phi^2 - dt^2}{\rho^2} - 2\rho dz(dt + d\phi) + \frac{1}{2}\rho^4(dt + d\phi)^2. \quad (2.3.14)$$

The limit $\rho \rightarrow 0$ in Poincaré coordinates corresponds to the timelike boundary \mathcal{I} of anti-de Sitter space. After setting $\tilde{\rho} = \rho^2$ the line element is

$$ds^2 = \frac{d\tilde{\rho}^2}{4\tilde{\rho}^2} + \frac{1}{\tilde{\rho}} \left(dz^2 + d\phi^2 - dt^2 - 2\tilde{\rho}^{3/2} dz(dt + d\phi) + \frac{1}{2}\tilde{\rho}^3(dt + d\phi)^2 \right), \quad (2.3.15)$$

and (2.3.15) is an expansion of the form

$$ds^2 = \frac{d\tilde{\rho}^2}{4\tilde{\rho}^2} + \frac{1}{\tilde{\rho}} g_{\mu\nu} dx^\mu dx^\nu, \quad g_{\mu\nu}(x, \tilde{\rho}) = g_{\mu\nu}^{(0)}(x) + g_{\mu\nu}^{(2)}(x)\tilde{\rho} + g_{\mu\nu}^{(3)}(x)\tilde{\rho}^{3/2} + \dots \quad (2.3.16)$$

as given in [80]. The coefficient $g^{(3)} = -dz \otimes (dt + d\phi) - (dt + d\phi) \otimes dz$ encodes the stress energy tensor of the boundary dual theory in the context of the AdS/CFT correspondence [15, 80]. It does not satisfy even the null energy condition on the three-dimensional conformal boundary, since $g_{\mu\nu}^{(3)} n^\mu n^\nu = -2$ for the null vector $n = \frac{\partial}{\partial t} + \frac{\partial}{\partial z}$. This presumably reflects causal pathologies of the bulk spacetime that we will try to understand by more conventional means shortly.

A general analysis of stationary cylindrically symmetric Einstein spaces was done in [109]. These authors assume the Lewis form of the metric, where a cross term $dz d\phi$ would be absent. We did not find it possible to bring (2.3.10) to the Lewis form. The theorem by Papapetrou [127] that any solution with two commuting Killing vectors (one timelike, one spacelike with periodic orbits) can be written in the Lewis form, only applies to solutions of the vacuum Einstein equations without cosmological term. Hence one cannot make a connection with spaces of the Lewis form as was possible for the Petrov solution.

2.3.2 Left-Invariant Forms

Since the action of the elements of the group manifold on itself has been given in (2.3.5), we can write down a matrix representation of this group of motions, with a

general element given by

$$g = \begin{pmatrix} 1 & 0 & 0 & 0 & r \\ 0 & e^r & 0 & 0 & z \\ 0 & 0 & e^{-\frac{r}{2}} \cos(\frac{\sqrt{3}}{2}r) & e^{-\frac{r}{2}} \sin(\frac{\sqrt{3}}{2}r) & \phi \\ 0 & 0 & -e^{-\frac{r}{2}} \sin(\frac{\sqrt{3}}{2}r) & e^{-\frac{r}{2}} \cos(\frac{\sqrt{3}}{2}r) & t \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \quad (2.3.17)$$

The group is generated by

$$Z = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \Phi = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (2.3.18)$$

$$T = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad R = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (2.3.19)$$

so that $g = e^{zZ + \phi\Phi + tT} e^{rR}$ and (r, z, ϕ, t) are coordinates on the group. The generators satisfy the algebra

$$[R, T] = -\frac{1}{2}T + \frac{\sqrt{3}}{2}\Phi, \quad [R, \Phi] = -\frac{1}{2}\Phi - \frac{\sqrt{3}}{2}T, \quad [R, Z] = Z. \quad (2.3.20)$$

This differs from the Killing algebra (2.3.3) by the usual overall minus sign coming from the fact that right-invariant vector fields generate left actions and vice versa.

The matrix representation (2.3.17) gives the group multiplication law

$$(r, z, \phi, t) \cdot (r', z', \phi', t') = (r + r', z + e^r z', \phi + e^{-\frac{r}{2}}(t' s + \phi' c), t + e^{-\frac{r}{2}}(t' c - \phi' s)), \quad (2.3.21)$$

where $s \equiv \sin \frac{\sqrt{3}}{2}r$, $c \equiv \cos \frac{\sqrt{3}}{2}r$. The Maurer-Cartan form is

$$g^{-1}dg = e^{-rR}(Z dz + \Phi d\phi + T dt)e^{rR} + R dr \equiv R \lambda^1 + Z \lambda^2 + \Phi \lambda^3 + T \lambda^4, \quad (2.3.22)$$

which gives the desired basis of left-invariant one-forms:

$$\begin{aligned} \lambda^1 &= dr, \quad \lambda^2 = e^{-r} dz, \quad \lambda^3 = e^{\frac{r}{2}} \left(\cos \left(\frac{\sqrt{3}}{2}r \right) d\phi - \sin \left(\frac{\sqrt{3}}{2}r \right) dt \right), \\ \lambda^4 &= e^{\frac{r}{2}} \left(\sin \left(\frac{\sqrt{3}}{2}r \right) d\phi + \cos \left(\frac{\sqrt{3}}{2}r \right) dt \right). \end{aligned} \quad (2.3.23)$$

We obtain a left-invariant metric on the group

$$\begin{aligned} ds^2 &= \eta_{ab} \lambda^a \otimes \lambda^b \\ &= dr^2 + e^{-2r} dz^2 + e^r \left(\cos(\sqrt{3}r)(d\phi^2 - dt^2) - 2 \sin(\sqrt{3}r) d\phi dt \right) \end{aligned} \quad (2.3.24)$$

with $\eta = \text{diag}(1, 1, 1, -1)$, which is the same as (2.3.1) and shows that our chosen coordinates agree with the initial Petrov coordinates. We see how to recover a metric on a group manifold; note that one could obtain this metric by just starting from the algebra (2.3.3).

2.3.3 Deformations of the Petrov Killing Algebra

We examine possible infinitesimal deformations of the four-dimensional Lie algebra. Using MATHEMATICA for the computations, equations (2.2.7) give the following conditions on linear deformations:

$$\begin{aligned} 0 &= A_z^r{}_\phi + \sqrt{3}A_t^r{}_z = A_z^r{}_t - \sqrt{3}A_\phi^r{}_z = A_\phi^r{}_t; \\ 0 &= 2A_\phi^r{}_r + A_\phi^z{}_z + \sqrt{3}A_t^z{}_z = -\sqrt{3}A_\phi^z{}_z + 2A_t^r{}_r + A_t^z{}_z = A_\phi^z{}_t = A_\phi^r{}_t; \\ 0 &= A_r^r{}_z + 2A_\phi^r{}_z + \sqrt{3}A_z^t{}_phi - \sqrt{3}A_t^phi{}_z = \sqrt{3}A_r^r{}_z + 2A_z^phi{}_t - \sqrt{3}A_z^t{}_t - \sqrt{3}A_\phi^phi{}_z \\ &= \sqrt{3}A_r^r{}_phi + A_t^phi{}_phi - \sqrt{3}A_\phi^t{}_t - A_t^r{}_r = \sqrt{3}A_z^r{}_phi - A_t^r{}_z; \\ 0 &= \sqrt{3}A_r^r{}_z - \sqrt{3}A_z^phi{}_phi - 2A_z^t{}_phi - \sqrt{3}A_t^t{}_z = A_r^r{}_z - \sqrt{3}A_z^phi{}_t + 2A_t^t{}_z + \sqrt{3}A_\phi^t{}_z \\ &= A_r^r{}_phi - \sqrt{3}A_\phi^phi{}_t + \sqrt{3}A_t^r{}_r - A_t^t{}_phi = A_z^r{}_phi + \sqrt{3}A_t^r{}_z. \end{aligned} \quad (2.3.25)$$

These constraints for $A_a{}^c{}_b$ reduce the number of free parameters from 24 to twelve. We list the most general deformation parameters satisfying (2.3.25) in a table:

Table 2.1: Infinitesimal deformations of the Petrov Killing algebra.

	$c = r$	$c = z$	$c = \phi$	$c = t$
$A_r^c{}_z$	$2C$	x_1	x_2	x_3
$A_r^c{}_phi$	$-\sqrt{3}A - B$	x_4	x_5	x_6
$A_r^c{}_t$	$-A + \sqrt{3}B$	x_7	x_8	x_9
$A_z^c{}_phi$	0	$2B$	C	$\sqrt{3}C$
$A_z^c{}_t$	0	$2A$	$-\sqrt{3}C$	C
$A_phi^c{}_t$	0	0	$-\sqrt{3}B - A$	$B - \sqrt{3}A$

The parameters $x_1, x_2, \dots, x_9, A, B$ and C can be arbitrary real constants. We need to investigate which of these correspond to trivial deformations. The conditions (2.2.9) mean that trivial deformations can be written as

$$A_r^r{}_z = -M^r{}_z, \quad A_r^z{}_z = -M^r{}_r, \quad A_r^\phi{}_z = -\frac{3}{2}M^\phi{}_z + \frac{\sqrt{3}}{2}M^t{}_z, \quad A_r^t{}_z = -\frac{\sqrt{3}}{2}M^\phi{}_z - \frac{3}{2}M^t{}_z. \quad (2.3.26)$$

The parameters C, x_1, x_2 and x_3 correspond to trivial deformations and can be set to zero by a change of basis. Furthermore,

$$A_r^r{}_\phi = \frac{1}{2}M^r{}_\phi + \frac{\sqrt{3}}{2}M^r{}_t, \quad A_r^r{}_t = -\frac{\sqrt{3}}{2}M^r{}_\phi + \frac{1}{2}M^r{}_t, \\ A_z^z{}_\phi = -M^r{}_\phi, \quad A_z^z{}_t = -M^r{}_t \quad (2.3.27)$$

etc., so that we can set $A = B = 0$,

$$A_r^z{}_\phi = \frac{3}{2}M^z{}_\phi + \frac{\sqrt{3}}{2}M^z{}_t, \quad A_r^\phi{}_\phi = -\frac{1}{2}M^r{}_r + \frac{\sqrt{3}}{2}M^\phi{}_t + \frac{\sqrt{3}}{2}M^t{}_\phi, \\ A_r^t{}_\phi = -\frac{\sqrt{3}}{2}M^r{}_r - \frac{\sqrt{3}}{2}M^\phi{}_\phi + \frac{\sqrt{3}}{2}M^t{}_t, \quad (2.3.28)$$

so that we can set $x_4 = x_5 = x_6 = 0$,

$$A_r^z{}_t = \frac{3}{2}M^z{}_t - \frac{\sqrt{3}}{2}M^z{}_\phi, \quad A_r^\phi{}_t = \frac{\sqrt{3}}{2}M^r{}_r - \frac{\sqrt{3}}{2}M^\phi{}_\phi + \frac{\sqrt{3}}{2}M^t{}_t, \\ A_r^t{}_t = -\frac{1}{2}M^r{}_r - \frac{\sqrt{3}}{2}M^\phi{}_t - \frac{\sqrt{3}}{2}M^t{}_\phi, \quad (2.3.29)$$

so that we can set $x_7 = 0$, but must treat x_8 and x_9 as nontrivial perturbations. After a relabelling of these parameters, the modified Lie algebra is now

$$[R, T] = -aT - b\Phi, \quad [R, \Phi] = -\frac{1}{2}\Phi - \frac{\sqrt{3}}{2}T, \quad [R, Z] = Z. \quad (2.3.30)$$

These relations satisfy the full Jacobi identity and so the linear deformation indeed defines a deformation of the Lie algebra. We may modify the matrix representation by setting

$$R = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & -b & 0 \\ 0 & 0 & -\frac{\sqrt{3}}{2} & -a & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (2.3.31)$$

In the case where $a \neq \frac{1}{2}$, one can always find a linear transformation of the basis vectors Φ and T such that the algebra takes the more symmetric form

$$[R, T] = a'T + b'\Phi, \quad [R, \Phi] = a'\Phi \pm b'T, \quad [R, Z] = Z, \quad (2.3.32)$$

with \pm depending on the value of b in the original deformation. This means that there are three distinct cases:

First case: Positive sign. Then a matrix representation of R is

$$R = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & a' & b' & 0 \\ 0 & 0 & b' & a' & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (2.3.33)$$

and a general group element looks like

$$g = e^{zZ + \phi\Phi + tT} e^{rR} = \begin{pmatrix} 1 & 0 & 0 & 0 & r \\ 0 & e^r & 0 & 0 & z \\ 0 & 0 & e^{a'r} \cosh(b'r) & e^{a'r} \sinh(b'r) & \phi \\ 0 & 0 & e^{a'r} \sinh(b'r) & e^{a'r} \cosh(b'r) & t \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \quad (2.3.34)$$

The Maurer-Cartan form is

$$\begin{aligned} g^{-1}dg &= Z e^{-r} dz + \Phi e^{-ar} (\cosh(b'r) d\phi - \sinh(b'r) dt) \\ &\quad + T e^{-ar} (\cosh(b'r) dt - \sinh(b'r) d\phi) + R dr \end{aligned} \quad (2.3.35)$$

and we can read off the left-invariant forms.

The special case $a' = 1$, $b' = 3$ gives the algebra of the Kaigorodov-Ozsváth solution (2.3.10), as can be seen by setting $\Phi' = \Phi - T$ and $T' = \Phi + T$, which amounts to

$$[R, \Phi] = \Phi + 3T, \quad [R, T] = 3\Phi + T \quad \Rightarrow \quad [R, \Phi'] = -2\Phi', \quad [R, T'] = 4T', \quad (2.3.36)$$

which is just the sign-reversed version of (2.3.13). One can recover the metric (2.3.10) by choosing the symmetric matrix

$$h_{ab} = \begin{pmatrix} -\frac{3}{\Lambda} & 0 & 0 & 0 \\ 0 & 1 & -\sqrt{2} & \sqrt{2} \\ 0 & -\sqrt{2} & -1 & -1 \\ 0 & \sqrt{2} & -1 & 3 \end{pmatrix} \quad (2.3.37)$$

and computing the left-invariant metric

$$h_{ab}\lambda^a \otimes \lambda^b = -\frac{3}{\Lambda}dr^2 + e^{-2r}dz^2 - 2e^{-2r}(d\phi^2 - dt^2) + e^{4r}(d\phi - dt)^2 - 2\sqrt{2}e^r dz(d\phi - dt), \quad (2.3.38)$$

which after the coordinate transformations $\phi - t = \phi'$ and $-\phi - t = t'$ reduces to (2.3.10). Since the matrix (2.3.37) has one negative and three positive eigenvalues, there exists a (vierbein) basis of left-invariant one-forms σ^a such that $\eta_{ab}\sigma^a \otimes \sigma^b$ gives the metric (2.3.10).

Second case: Negative sign.

$$R = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & a' & b' & 0 \\ 0 & 0 & -b' & a' & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad g = e^{zZ+\phi\Phi+tT}e^{rR} = \begin{pmatrix} 1 & 0 & 0 & 0 & r \\ 0 & e^r & 0 & 0 & z \\ 0 & 0 & e^{a'r}\cos(b'r) & e^{a'r}\sin(b'r) & \phi \\ 0 & 0 & -e^{a'r}\sin(b'r) & e^{a'r}\cos(b'r) & t \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.3.39)$$

The Maurer-Cartan form is

$$g^{-1}dg = Z e^{-r}dz + \Phi e^{-ar}(\cos(b'r)d\phi - \sin(b'r)dt) + T e^{-ar}(\cos(b'r)dt + \sin(b'r)d\phi) + R dr \quad (2.3.40)$$

and a left-invariant metric will be given by

$$ds^2 = \eta_{ab}\lambda^a \otimes \lambda^b = dr^2 + e^{-2r}dz^2 + e^{-2a'r}(\cos(2b'r)(d\phi^2 - dt^2) - 2\sin(2b'r)d\phi dt). \quad (2.3.41)$$

The original Petrov algebra is of course the special case $a' = -\frac{1}{2}$, $b' = \frac{\sqrt{3}}{2}$.

For the metric (2.3.41) the Ricci tensor has non-vanishing components

$$\begin{aligned} R_{rr} &= -1 - 2a'^2 + 2b'^2, \quad R_{zz} = -(1 + 2a')e^{-2r}, \\ R_{\phi\phi} &= -(1 + 2a')e^{-2a'r}(a'\cos(2b'r) + b'\sin(2b'r)), \\ R_{\phi t} &= (1 + 2a')e^{-2a'r}(a'\sin(2b'r) - b'\cos(2b'r)), \\ R_{tt} &= (1 + 2a')e^{-2a'r}(a'\cos(2b'r) + b'\sin(2b'r)). \end{aligned} \quad (2.3.42)$$

The manifold is an Einstein manifold only if $a' = -\frac{1}{2}$ and $b' = \pm\frac{\sqrt{3}}{2}$ ($\Lambda = 0$, Petrov solution) or $a' = 1$ and $b' = 0$ ($\Lambda = -3$, anti-de Sitter space). In the general case the energy-momentum tensor defined by $T_{\mu\nu} = \frac{1}{8\pi G}G_{\mu\nu}$ does not satisfy the weak energy condition; without loss of generality assume $\sin(\sqrt{3}r) = 0$ and $\cos(\sqrt{3}r) = 1$ and choose a timelike vector $t^\mu = (0, 0, t_3, t_4)$ ($t_4^2 \geq t_3^2$), then

$$G_{\mu\nu}t^\mu t^\nu = (1 + a' + a'^2 - b'^2)(t_3^2 - t_4^2) - 2(1 + 2a')b't_3t_4 \quad (2.3.43)$$

can be made arbitrarily negative by letting $t_3, t_4 \rightarrow \pm\infty$ while keeping $t_4^2 - t_3^2$ small and positive, unless $a' = -\frac{1}{2}$ and $b'^2 \geq \frac{3}{4}$. In the case $a' = -\frac{1}{2}$ and $b'^2 \geq \frac{3}{4}$, the Einstein tensor can be written as

$$G_{\mu\nu} = -\lambda g_{\mu\nu} + 2\lambda u_\mu u_\nu, \quad \lambda \equiv b'^2 - \frac{3}{4} \geq 0, \quad u_\mu = (1, 0, 0, 0), \quad u_\mu u^\mu = 1. \quad (2.3.44)$$

Note that u_μ is spacelike. This tensor satisfies the dominant energy condition as $2t^r v^r - t^a v_a \geq 0$ for any timelike and future-directed t, v . These statements are independent of the choice of the arrow of time, *i.e.* hold for both $t_4 < 0$ or $t_4 > 0$.

Third case: $a = \frac{1}{2}$. Let us introduce a new parameter c , so that

$$[R, T] = -\frac{1}{2}T - \frac{2}{\sqrt{3}}c^2\Phi, \quad [R, \Phi] = -\frac{1}{2}\Phi - \frac{\sqrt{3}}{2}T, \quad [R, Z] = Z \quad (2.3.45)$$

for positive b which gives

$$R = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & -\frac{2c^2}{\sqrt{3}} & 0 \\ 0 & 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (2.3.46)$$

and a general group element looks like

$$g = e^{zZ + \phi\Phi + tT} e^{rR} = \begin{pmatrix} 1 & 0 & 0 & 0 & r \\ 0 & e^r & 0 & 0 & z \\ 0 & 0 & e^{-\frac{r}{2}} \cosh(cr) & -\frac{2c}{\sqrt{3}} e^{-\frac{r}{2}} \sinh(cr) & \phi \\ 0 & 0 & -\frac{\sqrt{3}}{2c} e^{-\frac{r}{2}} \sinh(cr) & e^{-\frac{r}{2}} \cosh(cr) & t \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \quad (2.3.47)$$

The Maurer-Cartan form is

$$\begin{aligned} g^{-1}dg &= R dr + Z e^{-r} dz + \Phi \left(e^{\frac{r}{2}} \cosh(cr) d\phi + \frac{2c}{\sqrt{3}} e^{\frac{r}{2}} \sinh(cr) dt \right) \\ &\quad + T \left(e^{\frac{r}{2}} \cosh(cr) dt + \frac{\sqrt{3}}{2c} e^{\frac{r}{2}} \sinh(cr) d\phi \right). \end{aligned} \quad (2.3.48)$$

In the limit $c \rightarrow 0$ or $b \rightarrow 0$ this becomes

$$g^{-1}dg = R dr + Z e^{-r} dz + \Phi e^{\frac{r}{2}} d\phi + T \left(e^{\frac{r}{2}} dt + \frac{\sqrt{3}}{2} e^{\frac{r}{2}} r d\phi \right). \quad (2.3.49)$$

The only remaining case, namely negative b , is

$$[R, T] = -\frac{1}{2}T + \frac{2}{\sqrt{3}}c^2\Phi, \quad [R, \Phi] = -\frac{1}{2}\Phi - \frac{\sqrt{3}}{2}T, \quad [R, Z] = -Z, \quad (2.3.50)$$

which turns the hyperbolic into trigonometric functions.

If $c \neq 0$ one can rescale the coordinate t (for instance) and recover the same left-invariant forms as before, hence this does not give anything new. In the case $c = 0$ a left-invariant metric is given by

$$ds^2 = dr^2 + e^{-2r}dz^2 + e^r \left(\left(1 - \frac{3}{4}r^2\right) d\phi^2 - \sqrt{3}r d\phi dt - dt^2 \right). \quad (2.3.51)$$

The Einstein tensor for this metric can be written as

$$G_{\mu\nu} = \frac{3}{16}g_{\mu\nu} + \frac{3}{8}(\text{diag}(-2, e^{-2r}, 2e^r, 0))_{\mu\nu}. \quad (2.3.52)$$

This does not satisfy the weak energy condition as $G_{\mu\nu}t^\mu t^\nu = -\frac{3}{16}(3 + e^r) < 0$ for the timelike vector $t^\mu = (1, 0, 0, 1)$.

Of course, for all three possible cases considered, one could take any constant symmetric matrix with one negative and three positive eigenvalues instead of η_{ab} to construct a left-invariant metric; this gives a 10-parameter family (subject to constraints) of metrics whose Killing algebras contain one of the deformations of the four-dimensional Petrov Killing algebra as a subalgebra. One of these parameters can be absorbed into an overall scale, and any cross terms $dr dz$, $dr d\phi$ or $dr dt$ can be absorbed into a redefinition of z , ϕ and t . One is left with a 6-parameter family of metrics, by far most of which will not appear to have a physical interpretation. We have chosen to just give examples; more efficient methods of constructing general metrics with a given isometry group are certainly known.

2.3.4 Spinning Cylinders

In the case of an infinite rigidly rotating dust cylinder one could imagine trying to speed up this cylinder by shooting in particles with some angular momentum which enter the interior region on causal curves and increase the angular velocity a , so as to reach and surpass the critical value $aR = \frac{1}{2}$ above which CTCs appear. We will show that this is not possible.

The interior part of the general van Stockum solution

$$ds^2 = H(\rho)(d\rho^2 + dz^2) + L(\rho)d\phi^2 + 2M(\rho)d\phi dt - F(\rho)dt^2, \quad (2.3.53)$$

describing the region $\rho < R$ is given by [157]

$$H = \exp(-a^2 \rho^2), \quad L = \rho^2(1 - a^2 \rho^2), \quad M = a\rho^2, \quad F = 1. \quad (2.3.54)$$

As there are closed timelike curves for $\rho > \frac{1}{a}$ we require $aR \leq 1$. The exterior solution is for $a < \frac{1}{2}$, from now on setting $R = 1$ for simplicity which is no loss of generality,

$$H = e^{-a^2 \rho^{-2a^2}}, \quad L = \frac{\rho \sinh(3\epsilon + \theta)}{2 \sinh 2\epsilon \cosh \epsilon}, \quad M = \frac{\rho \sinh(\epsilon + \theta)}{\sinh 2\epsilon}, \quad F = \frac{\rho \sinh(\epsilon - \theta)}{\sinh \epsilon} \quad (2.3.55)$$

where $\theta(\rho) = \sqrt{1 - 4a^2} \log \rho$ and $\epsilon = \text{Artanh} \sqrt{1 - 4a^2}$. Note that always $-FL - M^2 = -\rho^2$ and so the metric has the right signature for all ρ (this of course is also true as $a \rightarrow \frac{1}{2}$). The point-particle Lagrangian is (a dot denotes differentiation with respect to an affine parameter λ)

$$\mathcal{L} = g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu = H(\dot{\rho}^2 + \dot{z}^2) + L\dot{\phi}^2 + 2M\dot{\phi}\dot{t} - F\dot{t}^2, \quad (2.3.56)$$

and since the Lagrangian does not depend on z, ϕ and t there are three conserved quantities associated with geodesics:

$$P \equiv H\dot{z}, \quad J \equiv L\dot{\phi} + M\dot{t}, \quad E \equiv F\dot{t} - M\dot{\phi}. \quad (2.3.57)$$

The Lagrangian for timelike or null geodesics becomes

$$\mathcal{L} = H\dot{\rho}^2 + \frac{P^2}{H} + \frac{1}{\rho^2} (FJ^2 - 2MEJ - LE^2) = -\mu^2 \leq 0 \quad (2.3.58)$$

and we obtain the radial equation

$$\begin{aligned} \left(\frac{d\rho}{d\lambda}\right)^2 &= \frac{L}{H\rho^2} \left(-\frac{\rho^2\mu^2}{L} - \frac{\rho^2 P^2}{HL} + E^2 + \frac{2M}{L}EJ - \frac{F}{L}J^2\right) \\ &= \frac{L}{H\rho^2} (E - V_{eff}^+(\rho))(E - V_{eff}^-(\rho)), \end{aligned} \quad (2.3.59)$$

where we have introduced an effective potential

$$V_{eff}^\pm(\rho) = \frac{M(\rho)}{L(\rho)}J \pm \rho \sqrt{\frac{1}{L(\rho)} \left(\mu^2 + \frac{P^2}{H(\rho)} + \frac{J^2}{L(\rho)}\right)}. \quad (2.3.60)$$

This is well-defined for all ρ as H, L, M all remain positive for all ρ . A particle falling in on a geodesic can enter the cylinder if

$$E > V_{eff}^+(1) = \frac{a}{1-a^2}J + \sqrt{\frac{1}{1-a^2} \left(\mu^2 + P^2 e^{a^2} + \frac{J^2}{1-a^2}\right)} \geq \frac{a+1}{1-a^2}J = \frac{1}{1-a}J. \quad (2.3.61)$$

Any particle entering the cylinder on a geodesic must have $\frac{J}{E} < 1 - a$. In the limit $a \rightarrow 0$ the conserved quantities J and E clearly describe angular momentum and energy per mass. We can identify the ratio $\frac{J}{E}$ with the angular velocity of an infalling particle at $R = 1$.

If we are considering accelerated observers, equation (2.3.58) still holds, but P , E and J will no longer be conserved quantities. However, only the local values of these quantities at $R = 1$ will decide about whether or not a particle will be able to enter the interior region of the cylinder.

This means that the above considerations also hold for accelerated observers and as any particle entering the cylinder must have $\frac{J}{E} < \frac{1}{2}$ for $a = \frac{1}{2}$, one cannot speed up the cylinder beyond $a = \frac{1}{2}$ using particles on timelike or null curves.

2.3.5 Geodesic Completeness

We ask whether the Petrov spacetime, with the radial coordinate r extended to take arbitrary values, is geodesically complete, *i.e.* whether all timelike and null geodesics can be extended to infinite values of the affine parameter. First we give an example that this need not be possible on a group manifold: Remove the null hyperplane $z = t$ from Minkowski space and consider the half-space $z > t$, denoted by M^- . It is clearly geodesically incomplete. Null translations and boosts

$$(z, t) \rightarrow (z + c, t + c); \quad (t + z, t - z) \rightarrow \left(\lambda(t + z), \frac{1}{\lambda}(t - z) \right) \quad (2.3.62)$$

act on M^- , and together with translations $(x, y) \rightarrow (x + a, y + b)$ they form a four-dimensional group which acts simply-transitively on M^- . For instance, the point $(x, y, z, t) = (0, 0, 1, 0)$ is, by a null translation and a successive boost, taken to

$$(0, 0, 1, 0) \rightarrow (0, 0, 1 + c, c) \rightarrow \left(0, 0, \lambda c + \frac{1}{2} - \frac{1}{2\lambda}, \lambda c + \frac{1}{2} + \frac{1}{2\lambda} \right). \quad (2.3.63)$$

There is a one-one correspondence between points in M^- and group parameters (a, b, λ, c) , where $\lambda > 0$. The space M^- can be identified with the group $G \times \mathbb{R}^2$, where G is the unique two-dimensional non-Abelian Lie group.

As a second example, introduced in a slightly different context in [76], consider the dilatation group generated by translations and dilatations

$$x^\mu \rightarrow x^\mu + c^\mu, \quad x^\mu \rightarrow \rho x^\mu, \quad (2.3.64)$$

where we denote the generators by P_a and D respectively. Parametrising the group elements by[‡] $g = e^{x^a P_a} e^{\lambda D}$, the Maurer-Cartan form is

$$g^{-1}dg = e^{-\lambda} dx^a P_a + d\lambda D \quad (2.3.65)$$

and hence a left-invariant metric is

$$ds^2 = d\lambda^2 + e^{-2\lambda} \eta_{\mu\nu} dx^\mu dx^\nu = \frac{1}{\rho^2} (d\rho^2 + \eta_{\mu\nu} dx^\mu dx^\nu), \quad (2.3.66)$$

where $\rho = e^\lambda$. This is the metric of anti-de Sitter space in five dimensions in Poincaré coordinates, which is geodesically incomplete as these coordinates cover only a patch of the full spacetime. Hence geodesic completeness is a non-trivial property of a group manifold.

To show geodesic completeness of the Petrov solution, we need to show that no geodesic reaches infinity for finite values of the affine parameter. Consider the Lagrangian

$$\mathcal{L} = g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu, \quad (2.3.67)$$

which for the metric (2.3.1) is

$$\mathcal{L} = \dot{r}^2 + e^{-2r} \dot{z}^2 + e^r \left(\cos \sqrt{3}r (\dot{\phi}^2 - \dot{t}^2) - 2\dot{\phi}\dot{t} \sin \sqrt{3}r \right). \quad (2.3.68)$$

Evidently, from the Euler-Lagrange equations, there are three conserved quantities because the Lagrangian does not depend on z , ϕ or t explicitly, the conjugate momenta

$$P \equiv e^{-2r} \dot{z}, \quad j \equiv e^r (\dot{\phi} \cos \sqrt{3}r - \dot{t} \sin \sqrt{3}r), \quad E \equiv e^r (\dot{\phi} \sin \sqrt{3}r + \dot{t} \cos \sqrt{3}r). \quad (2.3.69)$$

The Lagrangian now takes the form

$$\mathcal{L} = \dot{r}^2 + e^{2r} P^2 + e^{-r} \left(\cos \sqrt{3}r (j^2 - E^2) + 2Ej \sin \sqrt{3}r \right), \quad (2.3.70)$$

and since the Lagrangian is a conserved quantity in geodesic motion the equation

$$\left(\frac{dr}{d\lambda} \right)^2 = e^{-r} (E^2 \cos \sqrt{3}r - 2Ej \sin \sqrt{3}r - j^2 \cos \sqrt{3}r) - e^{2r} P^2 - \mu^2 \quad (2.3.71)$$

is satisfied by any geodesic, where μ^2 is positive for timelike, zero for null and negative for spacelike geodesics. For timelike geodesics ($\mu^2 > 0$), right-hand side of (2.3.71)

[‡]The coordinate λ only covers dilatations with $\rho > 0$, so that one may obtain geodesic completeness by adding the disconnected component with $\rho < 0$. This suggests that geodesic completeness of groups is a topological question that cannot be decided at the level of the Lie algebra.

becomes negative for large r , so that r is bounded and we may extend geodesics infinitely. For null geodesics, \dot{r} is bounded[§]. So for any finite values of the affine parameter, \dot{r} remains finite and so do $\dot{z}, \dot{\phi}$ and \dot{t} . Hence the Petrov spacetime is geodesically complete. It will be incomplete if we interpret it as the exterior solution of a rotating cylinder and cut off the region described by the Petrov solution at some value of r .

By very similar arguments we can show geodesic completeness of the Kaigorodov-Ozsváth solution with line element (2.3.14). In this case, the Lagrangian is

$$\mathcal{L} = \frac{1}{\rho^2} \left(\dot{\rho}^2 + \dot{z}^2 + \dot{\phi}^2 - \dot{t}^2 \right) - 2\rho \dot{z}(\dot{t} + \dot{\phi}) + \frac{1}{2}\rho^4(\dot{t} + \dot{\phi})^2 \quad (2.3.72)$$

and the conserved quantities are

$$P \equiv \frac{1}{\rho^2} \dot{z} - \rho(\dot{t} + \dot{\phi}), \quad j \equiv \frac{1}{\rho^2} \dot{\phi} - \rho \dot{z} + \frac{1}{2}\rho^4(\dot{t} + \dot{\phi}), \quad E \equiv \frac{1}{\rho^2} \dot{t} + \rho \dot{z} - \frac{1}{2}\rho^4(\dot{t} + \dot{\phi}) \quad (2.3.73)$$

so that the radial equation is

$$\left(\frac{d\rho}{d\lambda} \right)^2 = -\rho^4(P^2 + j^2 - E^2) - 2\rho^7 P(j + E) - \frac{1}{2}\rho^{10}(j + E)^2 - \rho^2 \mu^2. \quad (2.3.74)$$

The right-hand side becomes negative if $\mu^2 > 0$ for both $\rho \rightarrow \infty$ and $\rho \rightarrow 0$, hence all timelike geodesics are bounded in ρ . For $\mu = 0$ the right-hand side is either constant zero or becomes negative for large r , so null geodesics are bounded from above in ρ . For very small ρ , $\dot{\rho}$ goes to zero, so as before $\dot{\rho}$ is bounded for null geodesics. This shows geodesic completeness. Note that in the case of anti-de Sitter space (2.3.74) would be

$$\left(\frac{d\rho}{d\lambda} \right)^2 = -\rho^4(P^2 + j^2 - E^2) - \rho^2 \mu^2 \quad (2.3.75)$$

and depending on the magnitudes of E , j and P the right-hand side blows up as $\rho \rightarrow \infty$ for some geodesics, which can reach $\rho = \infty$ in finite affine parameter distance.

2.3.6 Causal Properties

It was observed by Tipler [157] that the Petrov solution contains CTCs. As shown in [158], its causal behaviour is actually even more pathological: The Petrov solution

[§]If we allow r to take negative values, then for $E \neq 0$ or $j \neq 0$ the right-hand side becomes oscillatory for large negative r , taking positive as well as negative values. Hence both timelike and null geodesics are bounded from below in r . For $E = j = 0$ the right-hand side is either constant zero or always negative.

is *totally vicious* (for a summary of causality conditions see [59]), *i.e.* for any two points a and b one has $a < b$ and $b < a$, where “ $<$ ” is the usual causal relation “can be joined by a timelike curve”. The proof of this statement proceeds by giving the future and past light cones at the identity element of the group manifold, and the observation that the adjoint action of elements of the form $\exp(\lambda R)$ rotates the cone around, such that for certain values of λ a future-directed timelike tangent vector is mapped to a past-directed timelike tangent vector.

One might hope that a similar statement could be made for the Kaigorodov-Ozsváth solution, but this turns out not to be case, as is apparent from the deformed group element (2.3.34) or the commutation relations (2.3.13). The action of R on the plane generated by Φ and T is now just a scaling of the generators, which amounts to an opening and closing of the light cones, but no rotation. The assumptions of Theorem 2 in [158] are not satisfied.

Apart from the results in [158] which refer to homogeneous spacetimes, we may try to apply a criterion for the causality properties of spacetimes with an Abelian isometry group acting transitively on timelike surfaces which was given by Carter in his 1967 PhD thesis [34]. For the Kaigorodov-Ozsváth solution, this Abelian isometry group is generated by $\frac{\partial}{\partial z}$, $\frac{\partial}{\partial \phi}$ and $\frac{\partial}{\partial t}$ which act on surfaces $\{r = \text{constant}\}$. Carter’s result is that a sufficient condition for such a spacetime to be totally vicious is the nonexistence of a one-form in the dual of the Abelian Killing algebra which is everywhere timelike or null. In the present case, such a one-form is of the form $\omega = A dz + B d\phi + C dt$, with A , B and C constants. Conversely, if there is such a one-form which is everywhere timelike, the spacetime is *virtuous*, *i.e.* causally well-behaved. In the present case, one finds that the norm of such an ω is

$$\tilde{g}(\omega, \omega) = e^{2r} \left(A^2 + 2\sqrt{2} A C e^{3r} + C(2B + C e^{6r}) \right), \quad (2.3.76)$$

where \tilde{g} is the contravariant version of (2.3.10). In order to make (2.3.76) non-positive for all values of r , one has to choose $A = C = 0$ to be left with the null one-form $d\phi$. Put differently, the subgroup generated by $\frac{\partial}{\partial z}$ and $\frac{\partial}{\partial t}$ acts on null surfaces $\{r = \phi = \text{constant}\}$; were these spacelike it would follow that the Kaigorodov-Ozsváth solution is causal (cf. Proposition 9 in [33]).

Carter’s condition for causality violation was used by Tipler in [157][¶] to show the pathological behaviour of the Petrov solution. Here the situation is less clear; we

[¶]whose wording is slightly unfortunate, since he states the relevant condition on ω as being “everywhere timelike”

cannot show directly that in our example spacetimes related by Lie algebra deformation have similar causal properties. The strongest indication we have for causal pathologies of the Kaigorodov-Ozsváth spacetime comes from the computation of the boundary stress-energy tensor in AdS/CFT in (2.3.16).

2.4 The Kaigorodov Solution and Lobatchevski Plane Gravitational Waves

As a different example of a homogeneous space which appears as an exact solution in general relativity, we consider the Kaigorodov solution, which has a well-known interpretation as the AdS analogue of homogeneous pp-waves on a Minkowski background [139]. It has received recent interest in the context of string theory as the near-horizon limit of an M2 brane [42, 129]. By using Lie algebra deformations we again try to find a spacetime which has a physical interpretation “close to” the Kaigorodov solution. An obvious candidate would be a member of the family of “Lobatchevski plane gravitational waves” discussed by Siklos in [149] as spacetimes admitting a Killing spinor. The expectation that the particular solution considered by Siklos which has a five-dimensional isometry group arises as a deformation of the Kaigorodov solution will be confirmed. Since the isometry group is now five-dimensional and not four-dimensional as in the previous section, one has to consider a quotient G/H . The larger symmetry will place constraints on the general form of a left-invariant metric.

2.4.1 The Kaigorodov Solution

The **Kaigorodov** solution (of Petrov type N) [93] is another example of a homogeneous space-time which solves the vacuum Einstein equations with a cosmological constant $\Lambda < 0$. The form given in [154] is

$$ds^2 = -\frac{12}{\Lambda}dz^2 + 10k e^{2z}dx^2 + e^{-4z}dy^2 - 10u e^z dz dx - 2e^z du dx. \quad (2.4.1)$$

The five-dimensional Killing algebra of (2.4.1) is spanned by

$$\begin{aligned} Y &\equiv \frac{\partial}{\partial y}, & X &\equiv \frac{\partial}{\partial x}, & U &\equiv e^{-5z} \frac{\partial}{\partial u}, \\ Z &\equiv \frac{\partial}{\partial z} + 2y \frac{\partial}{\partial y} - x \frac{\partial}{\partial x}, & W &\equiv ye^{-5z} \frac{\partial}{\partial u} + x \frac{\partial}{\partial y}. \end{aligned} \quad (2.4.2)$$

The stabiliser of the origin is the one-parameter subgroup generated by W , which we call H to be consistent with previous notation. The transformations generated by Z are

$$z \rightarrow z + \lambda, \quad y \rightarrow ye^{2\lambda}, \quad x \rightarrow xe^{-\lambda}, \quad (2.4.3)$$

while W generates

$$y \rightarrow y + \lambda x, \quad u \rightarrow u + ye^{-5z}\lambda + \frac{1}{2}xe^{-5z}\lambda^2. \quad (2.4.4)$$

The Killing vectors have non-zero commutators

$$\begin{aligned} [W, Y] &= -U, \quad [W, X] = -Y, \quad [Z, U] = -5U, \\ [Z, Y] &= -2Y, \quad [Z, X] = X, \quad [W, Z] = 3W; \end{aligned} \quad (2.4.5)$$

Y, X and U span a three-dimensional Abelian subalgebra. The generators of the kinematical group will have the sign-reversed algebra

$$\begin{aligned} [W, Y] &= U, \quad [W, X] = Y, \quad [Z, U] = 5U, \\ [Z, Y] &= 2Y, \quad [Z, X] = -X, \quad [W, Z] = -3W. \end{aligned} \quad (2.4.6)$$

We denote the isometry group by \mathcal{K}^{\parallel} ; it has three-dimensional subgroups of Bianchi types I, II and VI_h . The spacetime is naturally represented as a coset space

$$\mathcal{M} = \mathcal{K}/H. \quad (2.4.7)$$

Represented as matrices, the group \mathcal{K} is generated by

$$\begin{aligned} X &= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \\ U &= \begin{pmatrix} 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad W = \begin{pmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned} \quad (2.4.8)$$

^{||}The \mathcal{K} stands for Kaigorodov or Killing, according to taste

Any element of the group can be parametrised by

$$g = e^{xX+yY+ue^{5z}U} e^{zZ} e^{wW}, \quad (2.4.9)$$

such that (x, y, z, u) are coordinates on the coset space \mathcal{M} . The Maurer-Cartan form

$$\omega_{\text{MC}} \equiv g^{-1}dg = e_a \lambda^a, \quad (2.4.10)$$

where e_a are the group generators, gives a basis of left-invariant one-forms on \mathcal{K} :

$$\begin{aligned} \lambda^1 &= e^z dx, & \lambda^2 &= e^{-2z} dy - w e^z dx, & \lambda^3 &= dz, \\ \lambda^4 &= du - w e^{-2z} dy + \frac{1}{2} w^2 e^z dx + 5u dz, & \lambda^5 &= dw + 3w dz. \end{aligned} \quad (2.4.11)$$

As usual, one can now construct a left-invariant metric on the group by

$$ds^2 = G_{ab} \lambda^a \lambda^b, \quad (2.4.12)$$

where G_{ab} is constant and non-degenerate. Any such metric will depend on the coordinate w on W , since the Lie algebra splitting

$$\mathfrak{k} = \mathfrak{m} \oplus \mathfrak{h} \quad (2.4.13)$$

is not reductive: $[W, Z] = 3W$ where one would require $[\mathfrak{h}, \mathfrak{m}] \subset \mathfrak{m}$. (Recall the definition of a reductive geometry given in (1.1.6)).

The original Kaigorodov metric can be reconstructed by taking

$$\begin{aligned} ds^2 &= 10k(\lambda^1)^2 + (\lambda^2)^2 - \frac{12}{\Lambda}(\lambda^3)^2 - 2\lambda^1\lambda^4 + (\lambda^5)^2 \\ &= 10k e^{2z} dx^2 + e^{-4z} dy^2 - \frac{12}{\Lambda} dz^2 - 2e^z dx(du + 5u dz) + (dw + 3w dz)^2 \end{aligned} \quad (2.4.14)$$

and reducing to the coset by dropping the last term.

We summarise the most important physical properties of this solution, as discussed in more detail in [139]: The Kaigorodov spacetime has a curvature singularity at $z = \infty$, and $z = -\infty$ represents null and spacelike infinity; the metric approaches the AdS metric asymptotically as $z \rightarrow -\infty$. The patch covered by the coordinates (x, y, z, u) can be extended by changing coordinates to $q = e^{2z}$, and extending to negative q . The enlarged spacetime splits into two disjoint regions $q > 0$ and $q < 0$ which cannot mutually communicate. Then $X = \frac{\partial}{\partial x}$ is a timelike Killing vector for $q < 0$, and the enlarged spacetime is stationary in this region. This means that gravitational radiation can exist in a stationary spacetime.

2.4.2 Deformations of the Lie Algebra

In order to find infinitesimal deformations of the Killing algebra (2.4.6), we again consider the Jacobi identity at linear order (2.2.7), again using MATHEMATICA for explicit computations. Imposing these constraints on $A_a^c{}_b$, one is left with 20 free parameters, as shown in table 2.2.

Table 2.2: Infinitesimal deformations of the Kaigorodov Killing algebra.

	$c = x$	$c = y$	$c = z$	$c = u$	$c = w$
$A_x^c{}_y$	$(2\epsilon + \eta)/6$	$(2\delta + 5\tau)/3$	0	$-\gamma/4$	0
$A_x^c{}_z$	$-\omega + 2P$	$3\alpha + \beta$	$-(\delta + \tau)/3$	c_2	γ
$A_x^c{}_u$	$\sigma/3$	$\eta/2$	0	$5/3(\delta + \tau)$	0
$A_x^c{}_w$	$(c_9 - 2c_8 - 5c_7)/5$	c_1	$(2\epsilon + \eta)/6$	α	δ
$A_y^c{}_z$	$(6c_9 - 7c_8 - 15c_7)/5$	$-\omega + P$	$(2\epsilon + \eta)/3$	β	τ
$A_y^c{}_u$	0	$-2\sigma/3$	0	$(5\epsilon + 4\eta)/3$	0
$A_y^c{}_w$	m	c_7	$\sigma/3$	c_5	ϵ
$A_z^c{}_u$	$-6m$	c_8	$-5\sigma/3$	ω	η
$A_z^c{}_w$	c_3	c_4	$(3c_9 - c_8)/5$	c_6	P
$A_u^c{}_w$	0	$-m$	0	c_9	σ

Again it turns out that almost all of the possible deformations should be considered trivial (*i.e.* correspond to a change of basis in the Lie algebra). Only a one-parameter deformation, which can be given by ω , cannot be obtained by a basis transformation. This infinitesimal deformation satisfies the full Jacobi identity, and we obtain the new algebra

$$\begin{aligned}
[W, Y] &= U, \quad [W, X] = Y, \quad [Z, U] = (5 + \alpha)U, \\
[Z, Y] &= (2 + \alpha)Y, \quad [Z, X] = (\alpha - 1)X, \quad [W, Z] = -3W.
\end{aligned} \tag{2.4.15}$$

One would perhaps have expected that at least one possible deformation would be to deform the commutation relations of Z to arbitrary numbers:

$$[Z, U] = aU, \quad [Z, Y] = bY, \quad [Z, X] = cX, \quad [W, Z] = dW. \tag{2.4.16}$$

The one-parameter deformation we have obtained is the most general deformation of this type: By a rescaling of Z one can always keep one of the four numbers a, b, c, d

fixed. Then the Jacobi identity puts two constraints on changing the coefficients in the three other commutators, leaving a single deformation parameter.

Put differently, one is free to replace the deformed algebra (2.4.15) by any one-parameter deformation of the action of Z which satisfies the Jacobi identity, e.g.

$$\begin{aligned} [W, Y] &= U, \quad [W, X] = Y, \quad [Z, U] = 5U, \\ [Z, Y] &= (2 + \alpha)Y, \quad [Z, X] = (2\alpha - 1)X, \quad [W, Z] = (\alpha - 3)W. \end{aligned} \quad (2.4.17)$$

We will use this last form in the following. In the matrix representation we only have to modify the matrix for Z :

$$Z_\alpha = \begin{pmatrix} 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 2 + \alpha & 0 & 0 \\ 0 & 0 & 0 & 2\alpha - 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (2.4.18)$$

2.4.3 Construction of the Manifold

We are trying to construct a metric on the quotient space

$$\mathcal{M}' = \mathcal{K}'/H, \quad (2.4.19)$$

where the stabiliser H is again generated by W . We use the same parametrisation as before:

$$g = e^{xX+yY+ue^{5z}U} e^{zZ_\alpha} e^{wW} \quad (2.4.20)$$

By computing the Maurer-Cartan form one finds the left-invariant forms are now

$$\begin{aligned} \lambda^1 &= e^{(1-2\alpha)z} dx, \quad \lambda^2 = e^{-(2+\alpha)z} dy - w e^{(1-2\alpha)z} dx, \quad \lambda^3 = dz, \\ \lambda^4 &= du - w e^{-(2+\alpha)z} dy + \frac{1}{2} w^2 e^{(1-2\alpha)z} dx + 5u dz, \quad \lambda^5 = dw + (3 - \alpha)w dz. \end{aligned} \quad (2.4.21)$$

Again, it is not possible to find a left-invariant metric

$$ds^2 = G_{ab} \lambda^a \lambda^b \quad (2.4.22)$$

which is independent of w . We can, however write any such metric in the form

$$ds^2 = G_{55} (dm + A_\mu dx^\mu)^2 + \tilde{G}_{\mu\nu} dx^\mu dx^\nu, \quad (2.4.23)$$

where $x^\mu = (x, y, z, u)$, and demand that $\tilde{G}_{\mu\nu}$, in order to be a metric on the coset space, be independent of w . This reduces the number of free parameters in G from 15 to nine, and the coset metric is (setting G_{55} , which determines the overall scale, to -1)

$$ds^2 = p_3 e^{z(2-4\alpha)} dx^2 + p_1 e^{-2z(2+\alpha)} dy^2 + p_2 dz^2 - 2e^{z(1-2\alpha)} dx(p_1 du + (5p_1 u - p_4)dz), \quad (2.4.24)$$

where we have defined

$$p_1 := G_{25}^2 - G_{22}G_{55}, \quad p_2 = G_{35}^2 - G_{33}G_{55}, \quad p_3 = G_{15}^2 - G_{11}G_{55}, \quad p_4 = G_{15}G_{35} - G_{13}G_{55}. \quad (2.4.25)$$

It is clear that by shifting u by a constant one can always achieve that $p_4 = 0$. This means the most general metric on the deformed coset space is

$$ds^2 = p_3 e^{z(2-4\alpha)} dx^2 + p_1 e^{-2z(2+\alpha)} dy^2 + p_2 dz^2 - 2p_1 e^{z(1-2\alpha)} dx(du + 5u dz). \quad (2.4.26)$$

One discovers that the determinant of (2.4.26) is

$$\det g = -e^{-2(1+3\alpha)z} p_1^3 p_2, \quad (2.4.27)$$

and one finds that for Lorentzian signature $(-+++)$ we must have $p_1 > 0$ and $p_2 > 0$. One can verify that the metric, as it should be, has five Killing vectors

$$\begin{aligned} Y &\equiv \frac{\partial}{\partial y}, \quad X \equiv \frac{\partial}{\partial x}, \quad U \equiv e^{-5z} \frac{\partial}{\partial u}, \\ Z_\alpha &\equiv \frac{\partial}{\partial z} + (2+\alpha)y \frac{\partial}{\partial y} + (2\alpha-1)x \frac{\partial}{\partial x}, \quad W \equiv ye^{-5z} \frac{\partial}{\partial u} + x \frac{\partial}{\partial y}, \end{aligned} \quad (2.4.28)$$

which satisfy the algebra

$$\begin{aligned} [W, Y] &= -U, \quad [W, X] = -Y, \quad [Z_\alpha, U] = -5U, \\ [Z_\alpha, Y] &= -(2+\alpha)Y, \quad [Z_\alpha, X] = (1-2\alpha)X, \quad [W, Z_\alpha] = (3-\alpha)W. \end{aligned} \quad (2.4.29)$$

One finds that the Ricci tensor of (2.4.26) is, in the coordinate basis,

$$R_{\mu\nu} = -\frac{3(2+\alpha)^2}{p_2} g_{\mu\nu} - \frac{5e^{(2-4\alpha)z} p_3 \alpha (\alpha-3)}{p_2} \delta_\mu^x \delta_\nu^x. \quad (2.4.30)$$

There is a cosmological term with

$$\Lambda = -\frac{3(2+\alpha)^2}{p_2}, \quad (2.4.31)$$

consistent with $p_2 = -12/\Lambda$ for the original Kaigorodov solution, and a traceless second term which describes null dust. In an orthonormal frame, given by

$$\begin{aligned} E^0 &= du + e^{z(1-2\alpha)} \left(-\frac{p_3}{2p_1} + \frac{25p_1 u^2}{2p_2} + \frac{p_1}{2} \right) dx, \\ E^1 &= du + e^{z(1-2\alpha)} \left(-\frac{p_3}{2p_1} + \frac{25p_1 u^2}{2p_2} - \frac{p_1}{2} \right) dx, \\ E^2 &= \sqrt{p_2} dz - \frac{5p_1}{\sqrt{p_2}} u e^{z(1-2\alpha)} dx, \\ E^3 &= \sqrt{p_1} e^{-z(2+\alpha)} dy, \end{aligned} \quad (2.4.32)$$

the nonvanishing components of T^{ab} are

$$T^{00} = T^{01} = T^{10} = T^{11} = \frac{5p_3\alpha(3-\alpha)}{p_2 p_1^2}. \quad (2.4.33)$$

The number (2.4.33) is chosen to be positive so that T^{ab} describes an electromagnetic field; then depending on the value of α the sign of p_3 has to be chosen appropriately.

Using deformations of the isometry group of the Kaigorodov solution we have constructed a spacetime which belongs to a more general class of “Lobatchevski plane gravitational waves” discussed in [149]. To bring the metric (2.4.26) into the standard form we apply the coordinate transformation

$$q = \frac{\sqrt{p_2}}{2+\alpha} e^{(2+\alpha)z}, \quad \tilde{y} = \sqrt{p_1} y, \quad \tilde{u} = -p_1 e^{5z} u. \quad (2.4.34)$$

In these coordinates the metric becomes

$$ds^2 = \frac{p_2}{(2+\alpha)^2 q^2} [dq^2 + d\tilde{y}^2 + 2dx d\tilde{u} + H(q)dx^2], \quad (2.4.35)$$

with

$$H(q) = p_3 p_2^{\frac{\alpha-3}{2+\alpha}} ((2+\alpha)q)^{\frac{6-2\alpha}{2+\alpha}}. \quad (2.4.36)$$

In the notation of [154], we have the metric form,

$$ds^2 = \frac{3}{|\Lambda|q^2} (dq^2 + d\tilde{y}^2 + 2dx d\tilde{u} \pm q^{2k} dx^2) \quad (2.4.37)$$

with $k = \frac{3-\alpha}{2+\alpha}$ and the sign corresponding to the sign of p_3 . The condition for positive energy given in [154] is equivalent to

$$p_3(3-\alpha)\alpha > 0, \quad (2.4.38)$$

which is consistent with the remarks below (2.4.33). The case $k = -1$, corresponding to Defrise’s pure radiation solution [44] with enhanced symmetry, is the only value

of k which does not correspond to any choice of α . Metrics of the form (2.4.37) are conformal to pp-waves and a more detailed physical interpretation is given in [149].

Similar to the previous case of the Petrov solution, using a deformation of the Killing algebra of a given spacetime we have constructed a solution with similar physical properties, namely in this case a pp-wave on an AdS background which now describes gravitational radiation as well as a null Maxwell field.

2.5 Summary

We have seen how the mathematical theory of Lie algebra deformations can relate different physical situations, in this case exact solutions of general relativity. The Petrov and Kaigorodov-Ozsváth solutions, the unique vacuum solutions without and with cosmological constant with simply-transitive four-dimensional isometry groups, were shown to be related by Lie algebra deformations; they were shown to be both geodesically complete. While the Petrov solution displays pathological causal behaviour as first observed by Tipler [157] (for the more general class of Lanczos-van Stockum solutions), we were not able to determine the causal properties of the Kaigorodov-Ozsváth solution. There is however strong indication that the latter shows causal pathologies since we found an unphysical energy-momentum tensor on the boundary in the context of AdS/CFT. We have to leave a more concrete study of the relation of these two properties to further work. Another interesting question that has been left open is whether the Kaigorodov-Ozsváth solution can be understood as a certain limit of an exterior solution of a matter distribution, perhaps representing a rotating cylinder as well.** More generally its physical interpretation which is so far largely unknown deserves further study.

In a second calculation, we analysed deformations of the Killing algebra of the Kaigorodov spacetime and saw that one is naturally led to a solution including an electromagnetic field. Although we have not constructed new exact solutions here, we have shown how this more algebraic approach to exact solutions of general relativity can give new insights on physical properties.

**A class of rotating cylinder solutions with negative cosmological constant has been studied in [27], who discuss the appearance of CTCs as well, but leave out certain limiting null cases. Our study strongly suggests that the Kaigorodov-Ozsváth solution, if related to the results of [27], would be a limiting case not considered in [27].

Chapter 3

Deformed General Relativity and Torsion

*“Seit die Mathematiker über die Relativitätstheorie
hergefallen sind, verstehe ich sie selbst nicht mehr.” [54]*

Albert Einstein

3.1 Introduction

In the previous chapter, we have discussed exact solutions of general relativity which are homogeneous spaces G/H . Since these are special cases with more symmetry than one would expect in a generic physical situation, it is natural to investigate a more general geometric description in which the symmetry of a given homogeneous space is only present infinitesimally.

The passage from special to general relativity can be understood as making the symmetries of a homogeneous space local: General relativity is formulated in terms of (pseudo-)Riemannian geometry, which makes the symmetries of Euclidean space, which can be thought of as a homogeneous space $E(p, q)/SO(p, q)$ *, local, and can thus be formulated on arbitrary manifolds. Physically, this is the incorporation of Einstein’s equivalence principle: At a given point, any spacetime looks to an inertial observer like Minkowski space. In the vier-/vielbein formalism of general relativity (originally due to Cartan as well as the Cartan connections we will use in this chapter), one chooses a basis of the tangent space at each point, and thus a “moving frame” of

*For p spacelike and q timelike dimensions. We will set $p = d - 1$ and $q = 1$ in the following.

vector fields, such that the metric in this frame is constant and diagonal. The frame field then encodes the metric information of the manifold.

In this chapter, we shall investigate a generalised formulation of general relativity in terms of **Cartan geometry** which incorporates the symmetries of an arbitrary given homogeneous space G/H on an infinitesimal level, thus providing an extension of both Felix Klein's Erlangen Programme [96] and (pseudo-)Riemannian geometry, as summarised in the following picture that we essentially take from [148]:

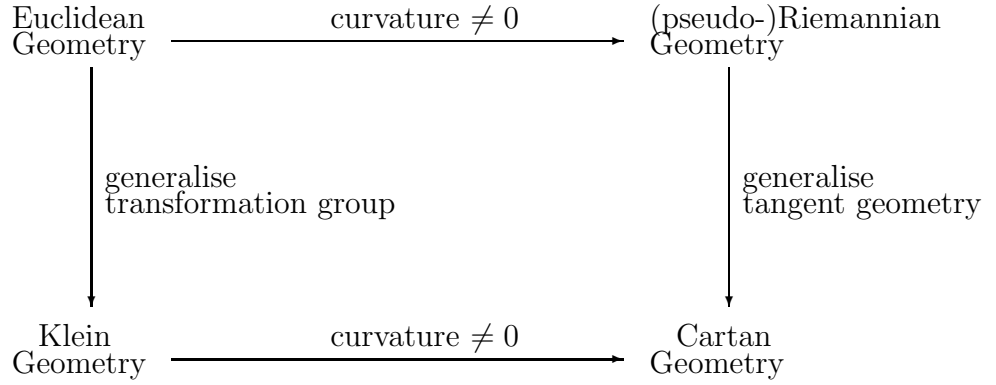


Figure 3.1: Cartan geometry extends both Klein geometry and (pseudo-)Riemannian geometry.

In general relativity, if an orthonormal frame has been found, one can still use general local Lorentz transformations to obtain a different frame with the same properties, and so in this formulation general relativity appears as a gauge theory. The description of general relativity is in terms of an \mathfrak{h} -valued Ehresmann connection (with $H = SO(d-1, 1)$, the Lorentz group), together with a soldering form (the frame field) determining how the tangent spaces at different points are “soldered” to the spacetime manifold. The gauge group H acts on the tangent space which is viewed as a vector space $\mathbb{R}^{d-1,1}$.

We will use a similar gauge-theoretic formulation, but of a more general character: Instead of viewing Minkowski space as $\mathbb{R}^{d-1,1}$, we will view it as a homogeneous space $E(d-1, 1)/SO(d-1, 1)$. A description in terms of Cartan geometry will allow us to *deform* general relativity by replacing $E(d-1, 1)$ by its deformation, the de Sitter group ${}^\dagger SO(d, 1)$. We effectively try to construct a theory of gravity built on tangent de Sitter spaces which replace the usual (co-)tangent spaces in general relativity. As

[†]The anti-de Sitter group $SO(d-1, 2)$ could be used equally well, and in fact is used in [153].

we saw in the previous chapter, from the viewpoint of Lie group deformations this process is a very natural extension of the standard framework. In addition to the speed of light c , a second scale κ with dimensions of momentum is introduced on a fundamental level. It is tempting to speculate that this scale could be related to quantum gravity, although the basic idea of “deforming” the kinematic algebra has existed at least since Heisenberg, who thought in terms of atomic rather than quantum-gravitational scales.

The motivation for investigating this possible generalisation of the usual formulation of Einstein-Cartan theory, or general relativity in first order formulation, comes from the recently proposed framework of deformed, or doubly, special relativity (DSR), where one effectively proposes a momentum space that is acted on by the de Sitter, not the Poincaré group, rather in the spirit of Snyder’s proposal [151]. It seems natural to try to construct a general geometric framework which incorporates this symmetry on the infinitesimal level. We will use a formulation of Einstein-Cartan theory given by Stelle and West, and interpret it in the concrete example of Minkowski space. By doing this construction, we hope to shed light on the physical interpretation of DSR, which seems rather unclear at present. In particular, while the noncommuting translations appearing in the de Sitter group are commonly interpreted as noncommuting “spacetime coordinates”, and one passes to a noncommutative geometry description to account for this, our discussion will stay within conventional differential geometry. We will see that the description of Minkowski space in this theory includes a connection with torsion instead of a fundamentally noncommutative structure.

In order to confront this mathematical framework with physical reality, we note that such a connection with torsion may lead to the usual ambiguities in minimal coupling. We give estimates concerning observable effects and note that observable violations of charge conservation induced by torsion should happen on a time scale of 10^3 s, which seems to rule out these modifications as a serious theory. Our considerations show, however, that the noncommutativity of translations in the Snyder algebra need not correspond to noncommutative spacetime in the usual sense.

To set the scene, we start with the Einstein-Hilbert action in four dimensions, without cosmological constant, written in first order (Palatini) form:

$$S_{\text{EH}} = \frac{1}{16\pi G} \int \epsilon_{abcd} (e^a \wedge e^b \wedge R^{cd}) , \quad (3.1.1)$$

where R^{cd} is the curvature two-form of an $\mathfrak{so}(3,1)$ -valued connection one-form ω^a_b .

One discovers that variation with respect to the vierbein field e^a gives the usual vacuum field equation $\text{Ric}_a = 0$ (where Ric_a is the Ricci one-form), while variation with respect to ω^a_b gives the condition of vanishing torsion. Thus, this formulation has the advantage over the second order form that the condition of a torsion-free connection does not have to be put in by hand. On the other hand, if one now adds a matter action to the Einstein-Hilbert action, both equations now potentially have a non-zero right-hand side. While the functional derivative of the matter action with respect to the vierbein determines an energy-momentum tensor, one will also have nonvanishing torsion if

$$\frac{\delta S_{\text{mat}}}{\delta \omega^a_b} \neq 0. \quad (3.1.2)$$

This is the case for matter with internal spin, such as spinor fields. This generalisation of general relativity which incorporates the possibility of torsion is known as **Einstein-Cartan theory**. Since torsion does not propagate, *i.e.* the field equation contains the full torsion two-form instead of contractions of it (as for curvature in Einstein's equations for $d > 3$), it seems difficult to rule out such a theory experimentally, since one would only measure torsion inside a matter distribution with spin.

The basic observation is now the following: Combine the connection ω^a_b and the vierbein e^a into a single connection, taking values in the algebra of the Poincaré group,

$$A = \begin{pmatrix} \omega^a_b & \frac{1}{L} e^i \\ 0 & 0 \end{pmatrix}, \quad (3.1.3)$$

where L is a length introduced to get dimensions right, but which is completely arbitrary at this stage. Then consider the Yang-Mills-like action (with now $d = 4$)

$$S = -\frac{L^2}{32\pi G} \int \epsilon_{abcd} (F^{ab} \wedge F^{cd}) = -\frac{L^2}{32\pi G} \int d^4x \frac{1}{4} \epsilon_{abcd} \epsilon^{\mu\nu\rho\tau} F_{\mu\nu}^{ab} F_{\rho\tau}^{cd}, \quad (3.1.4)$$

involving the curvature F^{ab} of the connection A^a_b (Latin indices here run from 1 to 4, and so one projects F^{ab} to its $\mathfrak{so}(3,1)$ part in this action). Writing out F^{ab} in terms of the curvature R^{ab} of ω^a_b and the vierbein field e^i , one discovers that the action (3.1.4) contains a topological Gauss-Bonnet term plus the Einstein-Hilbert action (3.1.1). Thus, the field equations of Einstein-Cartan theory are recovered.

The idea that general relativity may be interpreted as a gauge theory with the Poincaré group as gauge group is of course far from new. In the usual textbook

discussions, the significance of a gauge theory is expressed as the promotion of a global symmetry in a given theory to a local one by the introduction of an additional field which appears in gauge-covariant derivatives. This viewpoint was probably first elaborated in detail by Utiyama [159], who showed that the promotion of a global to a local symmetry directly leads to a gauge field, valued in the Lie algebra of the symmetry group under consideration, with the usual transformation properties, and thus geometrically to the well known formulation in terms of connections on principal bundles (see e.g. [120]). Utiyama also considered the Lorentz group as an example and was led to general relativity, but only by making several assumptions, such as the existence of a Riemannian metric and a local Lorentz (vierbein) frame and the symmetricity of the affine connection in its lower indices, without further motivation. It was then realised by Kibble [94] that the more natural gauge group to take is the Poincaré group, so that the vierbein appears naturally as part of the connection. Kibble showed that one is led to general relativity in the first order formalism, *i.e.* to Einstein-Cartan theory, and so to a connection with torsion when matter with internal spin, such as fermionic matter, is present. The appearance of torsion when fermions are coupled to gravity had earlier been noted by Weyl [165].[‡]

The result of MacDowell and Mansouri [110] is essentially that by considering a connection taking values in the algebra $\mathfrak{so}(4, 1)$ of the de Sitter group, one obtains a cosmological term in addition to the Einstein-Hilbert action (3.1.1). (MacDowell and Mansouri also extended the calculations to supergravity, which we will not discuss further.) We will see that the result is not just a mathematical trick, but has a geometric interpretation in terms of a Cartan connection.

The rest of this chapter is organised as follows: We discuss the idea of a curved momentum space in section 3.2 and give a brief introduction into the ideas of **deformed (doubly) special relativity** (DSR) most relevant to the following discussion in section 3.3. In section 3.4 we outline how Einstein-Cartan theory can be formulated as a gauge theory of gravity with the de Sitter group $SO(d, 1)$ as gauge group; this theory includes a gauge field that plays a crucial role in what follows. In this section we essentially re-derive the results of Stelle and West, using a different set of coordinates which we find more closely related to the DSR literature. This section uses the

[‡]Rather amusingly, Weyl wondered in [165] whether the “gauge transformations” of electromagnetism should not be called “phase transformations” instead – the term “gauge transformation” had of course been introduced by Weyl himself [163] in the context of local conformal transformations (changes in “gauge”).

language of gauge theory familiar to physicists; we then give a more mathematical account of the geometry of Cartan connections in section 3.5.

We then return to a concrete example: To justify our claim that this geometric framework can be used to generalise the ideas of DSR, we show in section 3.6 how, if spacetime is taken to be Minkowski space, the simplest non-trivial choice of zero section leads to a connection with torsion, providing a geometric interpretation for the noncommuting “coordinates” appearing in the Snyder algebra. We close with a discussion of our results and their possible physical implications, which show that the theory, at least in its given form, is not physically viable. We conclude that there may be different physical interpretations of algebraic commutation relations such as those used in DSR.

Since the two most obvious extensions of general relativity are admitting either connections with torsion or non-metric connections, we briefly discuss the theory of a torsion-free non-metric connection, known as symmetric affine theory, in a final section. We will see that it does not fit as well into a description by Cartan geometry as the case highlighted in this chapter.

3.2 Curved Momentum Space

It is commonly assumed that quantum gravity sets a fundamental length scale, the Planck scale [137], which can not be resolved by any physical experiment. Different approaches to quantum gravity, such as string theory or loop quantum gravity, incorporate such a scale. This leads to the idea that some kind of “space discreteness” should be apparent even in a low-energy “effective” theory.

The idea of putting quantum mechanics on a discrete lattice[§] seems to have been first considered by Heisenberg in the spring of 1930 [130], in an attempt to remove the divergence in the electron self-energy. Because the absence of continuous spacetime symmetries leads to violations of energy and momentum conservation, this approach was not pursued further, but later in the same year Heisenberg considered modifying the commutation relations involving position operators instead [130].

A fundamental length scale is absent in special relativity, where two observers will in general not agree on lengths or energies they measure. Hence the usual ideas of Poincaré invariance must be modified in some way. Snyder observed [151] that

[§]with spacing equal to the Compton wavelength of the proton, $l_c \approx 1.3$ fm

this could be done by deforming the Poincaré algebra into the de Sitter algebra, *i.e.* considering the isometry group of a (momentum) space of constant curvature. As we have seen in chapter 2, if one maintains the structure of a Lie algebra and considers deformations of the Poincaré algebra, replacing it by the de Sitter (or anti-de Sitter) algebra is the unique way of implementing a modified kinematic framework [12].

A d -dimensional de Sitter momentum space with curvature radius κ is defined as the submanifold of a $(d+1)$ -dimensional flat space with metric signature $(d, 1)$ by

$$(P^1)^2 + (P^2)^2 + \dots + (P^{d-1})^2 - (P^d)^2 + (P^{d+1})^2 = \kappa^2, \quad (3.2.1)$$

where κ has dimensions of mass. Its isometry group is generated by the algebra

$$\begin{aligned} [M_{ab}, M_{cd}] &= \eta_{ac}M_{bd} + \eta_{bd}M_{ac} - \eta_{bc}M_{ad} - \eta_{ad}M_{bc}, \\ [X_a, M_{bc}] &= \eta_{ac}X_b - \eta_{ab}X_c, \quad [X_a, X_b] = \frac{1}{\kappa^2}M_{ab}. \end{aligned} \quad (3.2.2)$$

Here M_{ab} correspond to a Lorentz subalgebra of the de Sitter algebra, while $X_a \equiv \frac{1}{\kappa}M_{d+1,a}$ are interpreted as (noncommuting) translations. These translations are then interpreted as corresponding to coordinates on spacetime; Snyder thought of operators acting on a Hilbert space. Since the operators X_1, \dots, X_{d-1} correspond to compact rotations in the $(d+1)$ -dimensional space, their spectrum is discrete. In this way, one obtains “quantised spacetime”, while maintaining Lorentz covariance.

It should be stressed at this point that we are interested in the full isometry group acting on momentum space, *i.e.* the Poincaré group in the case of a Minkowski space. Clearly, an equation such as (3.2.1) defines a scale κ already in the $(d+1)$ -dimensional Minkowski space; this scale would however not be invariant under translations of this space. In other words, there is no natural length scale in Minkowski space, whereas there is one in de Sitter space, namely the scale κ defining the manifold. There has been some confusion in the literature (see *e.g.* [147]) about the way in which Lorentz invariance is deformed in the Snyder algebra and the more recent ideas of DSR that we will review shortly. It should be clear from the commutation relations (3.2.2) that the Lorentz subgroup generated by M_{ab} is *not* deformed in any way.

One can give explicit expressions for the algebra elements by choosing coordinates on de Sitter space (3.2.1). The choice made by Snyder is taking Beltrami coordinates

$$p^1 = \kappa \frac{P^1}{P^{d+1}}, \quad p^2 = \kappa \frac{P^2}{P^{d+1}}, \dots, \quad p^d = \kappa \frac{P^d}{P^{d+1}}, \quad (3.2.3)$$

whence one has $(P^{d+1})^2 = \kappa^4/(\kappa^2 + \eta_{ab}p^ap^b)$ to satisfy (3.2.1), and $\eta_{ab}p^ap^b \geq -\kappa^2$, corresponding to an apparent maximal mass if p^a were interpreted as Cartesian co-

ordinates on a Minkowski momentum space. (Up to this point one could in principle have chosen anti-de Sitter instead of de Sitter space. Then this inequality becomes $\eta_{ab}p^ap^b \leq \kappa^2$, which perhaps seems less motivated physically.) A necessary sign choice means that these coordinates only cover half of de Sitter space. In these coordinates, the translation generators

$$X_a = \frac{1}{\kappa} \left(P^{d+1} \frac{\partial}{\partial P^a} - P_a \frac{\partial}{\partial P^{d+1}} \right) = \frac{\partial}{\partial p^a} + \frac{1}{\kappa^2} p_a p^b \frac{\partial}{\partial p^b} \quad (3.2.4)$$

generate “displacements” in de Sitter space. (In this notation, indices are raised and lowered with η_{ab} , the d -dimensional Minkowski metric, so that $p_a = \eta_{ab}p^b$.)

The motivation behind these ideas was to cure the infinities of quantum field theory, which evidently arise from allowing arbitrary high momenta (or short distances). In a somewhat similar spirit, Gol’fand suggested [73] to define quantum field theory on a momentum space of constant curvature, using Beltrami coordinates as momentum variables. This makes the volume of the corresponding Riemannian space finite and so presumably leads to convergent loop integrals in the Euclideanised theory. The consequences for standard quantum field theory were further explored in [92, 74].

Gol’fand only assumed that $\kappa \gg m$ for all elementary particles; thinking of quantum gravity, one would perhaps identify κ with the Planck scale, whereas the original authors seem to have thought of the Fermi (*i.e.* weak interaction) scale.

The induced metric on de Sitter space in terms of the coordinates p^a is

$$g_{nr} = \frac{\kappa^2}{\kappa^2 + p \cdot p} \left(\eta_{nr} - \frac{p_n p_r}{\kappa^2 + p \cdot p} \right), \quad (3.2.5)$$

where $p \cdot p \equiv \eta_{cd}p^c p^d$. The metric (3.2.5) becomes singular when $p \cdot p \rightarrow -\kappa^2$, and negative definite when extended to what Gol’fand calls the exterior region $p \cdot p < -\kappa^2$. In four dimensions,

$$\det g = -\kappa^{10} (\kappa^2 + p \cdot p)^{-5}, \quad (3.2.6)$$

and the volume element is $d^4p \kappa^5 (\kappa^2 + p \cdot p)^{-5/2}$.

In Gol’fand’s approach (assuming $d = 4$ of course), the standard Feynman rules were modified by replacing the addition of momenta p and k at a vertex by

$$(p(+))k^a = \frac{\kappa}{\kappa^2 - p \cdot k} \left(p^a \sqrt{\kappa^2 + k \cdot k} + k^a \left(\kappa - \frac{p \cdot k}{\kappa + \sqrt{\kappa^2 + k \cdot k}} \right) \right), \quad (3.2.7)$$

which corresponds to a translation by k of the vector p . (Again $p \cdot k \equiv \eta_{ab}p^a k^b$, etc.) It was also noted that spinors now transform under “displacements” as well, which is

made more explicit in [92] and [74]. As is well known, five-dimensional Dirac spinors still have four components and the matrix γ^5 appears in the Dirac Lagrangian, hence there is no chirality. This alone seems to imply that the original Gol’fand proposal cannot be used for an appropriate model of the known particles.

Gol’fand’s approach is very different from more recent approaches to quantum field theory on noncommutative spaces (see *e.g.* [48]) in that the field theory is defined on a momentum space which is curved, but neither position nor momentum space are noncommutative in the usual sense.

3.3 Deformed Special Relativity

The idea that the classical picture of Minkowski spacetime should be modified at small length scales or high energies was re-investigated in more recent times, motivated by the apparent existence of particles in ultra high energy cosmic rays whose energies could not be explained within special relativity [5]. The proposed framework of deformed special relativity (DSR) [4] modifies the Poincaré algebra, introducing an energy scale κ into the theory, in addition to the speed of light c . This leads to a quantum (κ -)deformation of the Poincaré algebra [113], with the parameter κ associated with the newly introduced scale.

It was soon realised [100] that this deformed algebra is the algebra of the isometry group of de Sitter space, and that the symmetries of DSR could hence be obtained by identifying momentum space with de Sitter space, identifying X_a as the generators of translations on this space. The constructions of DSR thus appear to be a resurrection of Snyder’s and Gol’fand’s ideas. We take this observation as the defining property of DSR, and will seek to describe a framework in which momentum space, or rather the (co-)tangent space in general relativity, is replaced by an “internal” de Sitter space. We will see that this can best be done using Cartan geometry.

When discussing DSR as a modification of special relativity, we take the view that special relativity is defined as a kinematic framework with preferred inertial systems, related to one another by (proper) Lorentz transformations. That is, one has a flat spacetime on which there exist certain preferred coordinate systems, those in which the metric is diagonal with entries ± 1 . From this point of view, the choice of coordinates on the internal de Sitter space plays quite an important role if one is looking for a “deformation” of special relativity including an energy scale κ . Such

a deformation can only arise if the chosen coordinate system reduces to Cartesian coordinates on Minkowski space as $\kappa \rightarrow \infty$. The choice of coordinates is obviously not unique.

The generators of the algebra will take different explicit forms when different coordinate systems (on four-dimensional de Sitter space) are chosen. In [100] “natural coordinates” are defined by, in the notation of section 3.2, ¶

$$g = \exp [p^I (M_{I4} + X_I)] \exp [p^4 X_4] \mathcal{O}, \quad (3.3.1)$$

where $\mathcal{O} = (0, 0, 0, 0, \kappa)$ is taken to be the origin of de Sitter space in five-dimensional Minkowski space, and M_{I5} and M_{45} correspond to translations in space and time. The coordinates one obtains are related to the five-dimensional coordinates by

$$P^I = p^I e^{\frac{p^4}{\kappa}}, \quad P^4 = \kappa \sinh \left(\frac{p^4}{\kappa} \right) + \frac{\vec{p}^2}{2\kappa} e^{\frac{p^4}{\kappa}}, \quad P^5 = \kappa \cosh \left(\frac{p^4}{\kappa} \right) - \frac{\vec{p}^2}{2\kappa} e^{\frac{p^4}{\kappa}}. \quad (3.3.2)$$

Again, these cover only half of de Sitter space where $P^4 + P^5 > 0$. The metric in these “flat” coordinates is

$$ds^2 = -(dp^4)^2 + e^{\frac{2p^4}{\kappa}} \delta_{IJ} dp^I dp^J. \quad (3.3.3)$$

Slices of constant p_4 are flat; to an observer using these coordinates the spacetime appears as expanding exponentially. An illuminating discussion of different coordinate systems and kinematics on de Sitter space is given in [29].

The Magueijo-Smolín model [112] corresponds to the following choice of coordinates:

$$p^1 = \kappa \frac{P^1}{P^5 - P^4}, \quad p^2 = \kappa \frac{P^2}{P^5 - P^4}, \quad p^3 = \kappa \frac{P^3}{P^5 - P^4}, \quad p^4 = \kappa \frac{P^4}{P^5 - P^4}, \quad (3.3.4)$$

The generators of boosts in de Sitter space take the form

$$K^I \equiv p^I \frac{\partial}{\partial p^4} + p^4 \frac{\partial}{\partial p^I} + \frac{1}{\kappa} p^I p^J \frac{\partial}{\partial p^J}, \quad (3.3.5)$$

and translations (not considered by the authors of [112]) would take the form

$$X_I = \frac{p^4 + \kappa}{\kappa} \frac{\partial}{\partial p^I} + \frac{1}{\kappa^2} p_I p^b \frac{\partial}{\partial p^b}, \quad X_4 = \frac{1}{\kappa} p^b \frac{\partial}{\partial p^b} + \frac{p^4 + \kappa}{\kappa} \frac{\partial}{\partial p^4}. \quad (3.3.6)$$

¶Capital Latin indices such as I and J used in this section only run over spatial coordinates (from 1 to 3).

This choice of coordinates is somewhat peculiar as p^4 takes a special role, as is also apparent from the modified dispersion relations presented in [112]. The quantity

$$||p||^2 = \frac{\eta_{ab} p^a p^b}{(1 + \frac{1}{\kappa} p^4)^2} \quad (3.3.7)$$

is invariant under boosts and rotations in de Sitter space, as would $\eta_{ab} p^a p^b$ be in Beltrami coordinates.

Each DSR model corresponds to a choice of coordinates on de Sitter space, such that all expressions reproduce the expressions for special-relativistic Minkowski coordinates as $\kappa \rightarrow \infty$. What Magueijo and Smolin call a “ U map” is essentially a coordinate transformation from Beltrami coordinates to a different set of coordinates, which becomes the identity as $\kappa \rightarrow \infty$. In the remaining sections we shall use Beltrami coordinates. Note that this means we always have $p \cdot p \geq -\kappa^2$.

3.4 A de Sitter Gauge Theory of Gravity

The most direct implementation of the ideas discussed so far into a framework describing more general spacetimes is replacing the cotangent (or tangent) bundle usually taken as phase space by a general symplectic manifold $\{\mathcal{P}, \omega\}$, which can be locally viewed as a product $U \times \mathcal{D}$ of a subset $U \subset \mathcal{M}$ of spacetime \mathcal{M} with de Sitter space \mathcal{D} . We want to retain the differentiable structure of a manifold, which we do not assume to be present in a full theory of quantum gravity. We also assume that the structure of momentum space is fixed and in particular does not depend on matter fields, as suggested in [119].

If phase space is described as such a manifold, with a choice of origin in the “tangent” de Sitter space at each point, the appropriate mathematical language is that of fibre bundles. The theory of connections on fibre bundles of this type, called *homogeneous bundles* in [51], was developed by Élie Cartan (*e.g.* in [32]). Adopting this framework means there is now an $\mathfrak{so}(d, 1)$ connection, instead of an $\mathfrak{so}(d - 1, 1)$ connection, defining parallel transport on spacetime.

It was noted by MacDowell and Mansouri [110] that gravity with a cosmological term in four dimensions could be obtained from a theory of such an $\mathfrak{so}(d, 1)$ connection by projecting it onto its $\mathfrak{so}(d - 1, 1)$ part in the action. A more elaborate description in terms of Einstein-Cartan theory was then given by Stelle and West [153]. Their analysis included the gauge field needed to identify the fibres at different spacetime

points, which will be crucial for the interpretation of the theory. The mathematical side of MacDowell-Mansouri gravity as a theory of a Cartan connection is nicely illustrated in [166]; we follow this article as well as the more computationally based presentation of [153], who use the language of non-linear realisations. An overview over the mathematics of Cartan connections is given in [148].

For clarity we first describe the framework in a language more common to physicists; a more mathematical account of Cartan connections on homogeneous bundles is given in the following section 3.5.

The usual description of general relativity as a gauge theory of the Lorentz group, known as vier-/vielbein formalism, was reviewed at the beginning of this chapter. Since the tangent bundle is in our description replaced by a homogeneous bundle with a curved “tangent” space, one has to effectively use a “double vielbein” formalism, in which spacetime vectors are mapped to vectors in the tangent space to the internal (curved) space by a soldering form (vielbein). The picture we have in mind is that of a de Sitter space rolled along the manifold. One then needs to introduce a new field which specifies the point of tangency, expressed in a given coordinate system on the internal space, at each spacetime point. We denote it by $p^a(x)$. This corresponds mathematically to a necessary choice of zero section (see below), and physically to a gauge field. Picking a point of tangency at each spacetime point breaks the gauge group $SO(d, 1)$ down to the Lorentz subgroup $SO(d - 1, 1)$ leaving the point of tangency invariant.

In more general terms [167], one has a principal bundle over a base manifold \mathcal{M} with gauge group G . Locally, a connection on this bundle can be viewed as a one-form taking values in the Lie algebra \mathfrak{g} . One then considers the projection of the connection to its $\mathfrak{g}/\mathfrak{h}$ part and demands that at any point in the manifold, this be an isomorphism between the tangent space at this point and the algebra $\mathfrak{g}/\mathfrak{h}$. In other words, the $\mathfrak{g}/\mathfrak{h}$ part e_μ^i , written as a matrix, has to have full rank. This of course means that the dimension of $\mathfrak{g}/\mathfrak{h}$ must be equal to the dimension of the manifold. The gauge field $p^a(x)$ now comes in because the projection of the connection to its $\mathfrak{g}/\mathfrak{h}$ part is not canonical; for a given subgroup $H \subset G$ viewed as the stabiliser of a particular point in G/H , one could equally well take a conjugate group $gHg^{-1} \simeq H$ stabilising a different point in G/H . This will become more explicit shortly when we consider the MacDowell-Mansouri action.

The procedure does not depend on the choice of gauge group G , and would be

equally possible if the gauge group were chosen to be the Poincaré group $E(d-1, 1)$; such a formulation for gravity, including coupling to matter, is indeed discussed in [75]. In that case, the homogeneous space G/H would be Minkowski space, viewed as an affine space. As the value of the gauge field $p^a(x)$ at a particular spacetime point picks an origin in this affine space, it may then be identified with the vector space $\mathbb{R}^{d-1,1}$: At a given point in Minkowski space, the tangent space to this point can be identified with Minkowski space itself. Although all calculations we will present go through in the case of $E(d-1, 1)$ as gauge group (resulting in different formulae of course), the formalism may seem a bit redundant since there is no real necessity for the second vielbein field. In the more general case of a homogeneous space G/H which acts as a “tangent space” instead of a flat vector space, it is however necessary.

We consider a theory with gauge group $SO(d, 1)$, so that the connection A takes values in the Lie algebra $\mathfrak{so}(d, 1)$. It can be split as (again introducing a length l on dimensional grounds)

$$A = \begin{pmatrix} \omega^a_b & \frac{1}{l}e^i \\ -\frac{1}{l}e_i & 0 \end{pmatrix}, \quad (3.4.1)$$

so that ω^a_b acts as the usual $\mathfrak{so}(d-1, 1)$ -valued connection of general relativity and e^i as a vielbein one-form. We have simultaneously unified the usual connection and the vielbein, and replaced the (flat) tangent space by a curved “internal” space, such that the de Sitter group and not the Poincaré group now appears as a gauge group. (Lorentz) indices on ω^a_b and e^i are raised and lowered using η^{ab} .

A gauge transformation, *i.e.* a local transformation $g(x)$ taking values in the de Sitter group, can be split as $g(x) = s(x)\Lambda(x)$, where $s(x)$ changes the zero section, *i.e.* changes the local identification of points of tangency at each spacetime point, and $\Lambda(x)$ is a local Lorentz transformation in the vielbein formalism of general relativity which does not mix the ω^a_b and e^i parts of the connection. In this notation, the connection transforms under a gauge transformation as

$$A(x) \rightarrow A'(x) = \Lambda^{-1}(x)s^{-1}(x)A(x)s(x)\Lambda(x) + \Lambda^{-1}(x)s^{-1}(x)ds(x)\Lambda(x) + \Lambda^{-1}(x)d\Lambda(x). \quad (3.4.2)$$

One can use this equation to relate the connection A_0 corresponding to the trivial zero section, where the point of tangency is the origin of the internal space at each spacetime point, $p^a(x) \equiv (0, 0, 0, 0)$, to a connection corresponding to any given

zero section. The physical significance of this is the following. Assume we have fixed $p^a(x) \equiv (0, 0, 0, 0)$. Then an action can be defined from the curvature of the connection A (here R is the curvature of the $\mathfrak{so}(d-1, 1)$ part of A),

$$F = dA + A \wedge A = \begin{pmatrix} R^a_b - \frac{1}{l^2}(e^a \wedge e_b) & \frac{1}{l}T^i \equiv \frac{1}{l}(de^i + \omega^i_j \wedge e^j) \\ -\frac{1}{l}T_i & 0 \end{pmatrix}. \quad (3.4.3)$$

In four dimensions, the MacDowell-Mansouri action [110, 166] is

$$S = -\frac{3}{32\pi G\Lambda} \int \epsilon_{abcd} (F^{ab} \wedge F^{cd}) = -\frac{3}{32\pi G\Lambda} \int d^4x \frac{1}{4} \epsilon_{abcd} \epsilon^{\mu\nu\rho\tau} F_{\mu\nu}^{ab} F_{\rho\tau}^{cd}, \quad (3.4.4)$$

where the Latin indices run from 1 to 4, and so one projects F to its $\mathfrak{so}(3, 1)$ part in this action.

Apart from a topological Gauss-Bonnet term, the action (3.4.4) is equivalent to the Einstein-Hilbert action with a cosmological term

$$S = \frac{3}{16\pi G\Lambda} \frac{1}{l^2} \int \epsilon_{abcd} \left(e^a \wedge e^b \wedge R^{cd} - \frac{1}{2l^2} e^a \wedge e^b \wedge e^c \wedge e^d \right), \quad (3.4.5)$$

where we have to identify

$$\Lambda = \frac{3}{l^2} \quad (3.4.6)$$

as the cosmological constant.^{||}

In order to define the projection of F in the action (3.4.4), one has used a splitting

$$\mathfrak{so}(d, 1) \simeq \mathfrak{so}(d, 1)/\mathfrak{so}(d-1, 1) \oplus \mathfrak{so}(d-1, 1), \quad (3.4.7)$$

which depends on the gauge field since the subgroup $SO(d-1, 1)$ leaving a given point in de Sitter space invariant depends on the choice of this point.

When the action (3.4.4) is coupled to matter, the $\mathfrak{so}(d, 1)/\mathfrak{so}(d-1, 1)$ part e^a of the connection appears in a volume element in the matter Lagrangian. By varying the action one obtains the field equations of Einstein-Cartan theory with a cosmological constant $\Lambda = 3/l^2$. The length scale l , which is so far arbitrary, can be chosen to reproduce the Λ of the observed universe, which means it must be chosen to be very large (the “cosmological constant problem”). By the field equations, one can

^{||}Note that in the formulation in terms of the Poincaré group, with connection (3.1.3), no cosmological constant appears and the length scale L remains undetermined.

determine for a given matter distribution a connection A_0 consisting of an $\mathfrak{so}(d-1, 1)$ connection $(\omega^a_b)_0$ and a vielbein e_0^i .

The MacDowell-Mansouri action reproducing Einstein-Cartan theory with a cosmological constant includes a gauge choice. We can hence view it as the gauge-fixed version of a more general theory. Since (3.4.2) determines how the connection transforms under a gauge transformation, we can generalise a given solution of Einstein-Cartan theory to an arbitrary gauge choice. The extension of the theory to arbitrary configurations of the gauge field, and hence arbitrary choices of tangency points of the internal space to spacetime, is what we call **Einstein-Cartan-Stelle-West theory**. Any solution of Einstein-Cartan theory, in particular any (torsion-free) solution of general relativity, gives rise to more general solutions of Einstein-Cartan-Stelle-West theory via (3.4.2). We will later see that one can construct an $\mathfrak{so}(d-1, 1)$ connection with torsion from a torsion-free one.

In (3.4.2), $s(x)$ takes values in the de Sitter group, more precisely in the subgroup generated by “translations” which leaves no point of de Sitter space invariant. The correspondence between Beltrami coordinates $p^a(x)$ on de Sitter space and such group elements is given explicitly by

$$s(p(x)) = \exp \left[\frac{p^i(x)}{\sqrt{-p(x) \cdot p(x)}} \operatorname{Artanh} \left(\frac{\sqrt{-p(x) \cdot p(x)}}{\kappa} \right) \kappa X_i \right]. \quad (3.4.8)$$

Then the group element $s(p(x))$ maps $(0, 0, 0, 0)$ to $(p^1(x), p^2(x), p^3(x), p^4(x))$ in Beltrami coordinates. A different choice of coordinates in the internal de Sitter space would correspond to a different parametrisation of the elements of the subgroup of translations of the de Sitter group.

Inserting (3.4.8) into (3.4.2) and setting $\Lambda(x) \equiv e$, we obtain

$$\begin{aligned} \omega^{ab}(p(x)) &= \frac{p^a e_0^b}{l\kappa\gamma(p)} + \left(1 - \frac{1}{\gamma(p)}\right) \frac{p^a dp^b + \omega_0^{ca} p^b p_c}{p \cdot p} + \frac{1}{2} \omega_0^{ab} - (a \leftrightarrow b), \\ e^i(p(x)) &= \frac{l\kappa}{p \cdot p + \kappa^2} \left(p^i \frac{p_c dp^c}{p \cdot p} (1 - \gamma(p)) + dp^i \gamma(p) + (\omega^i_b)_0 p^b \gamma(p) \right) \\ &\quad + p^i e_0^a p_a \frac{1 + \frac{\kappa^2}{p \cdot p} (1 - \gamma(p))}{p \cdot p + \kappa^2} + \frac{e_0^i}{\gamma(p)}, \end{aligned} \quad (3.4.9)$$

where

$$\gamma(p) \equiv \sqrt{\frac{p \cdot p + \kappa^2}{\kappa^2}} = 1 + \frac{p \cdot p}{2\kappa^2} + \dots \quad (3.4.10)$$

Because $p \cdot p \geq -\kappa^2$ in Beltrami coordinates, the square root is always real. In the limit $p \cdot p \rightarrow 0$, our parametrisation is the same as that used in [153], and we recover

their results

$$\begin{aligned}\omega^{ab}(p(x)) &= \left(\frac{1}{2}\omega_0^{ab} + \frac{1}{l\kappa}p^a e_0^b + \frac{1}{2\kappa^2}(p^a dp^b + \omega_0^{ca}p^b p_c) \right) - (a \leftrightarrow b), \\ e^i(p(x)) &= e_0^i + \frac{l}{\kappa} \left(-\frac{1}{2\kappa^2}p^i p_c dp^c + dp^i + \omega_0^{ib}p_b \right) + \frac{1}{2\kappa^2}p^i e_0^a p_a.\end{aligned}\quad (3.4.11)$$

Near $p = 0$, we have

$$\omega^{ab}(p(x)) = \omega_0^{ab} + O\left(\frac{p}{\kappa}\right), \quad e^i(p(x)) = e_0^i + \frac{l}{\kappa}dp^i + O\left(\frac{p}{\kappa}\right). \quad (3.4.12)$$

As mentioned above, the $\mathfrak{so}(d, 1)/\mathfrak{so}(d - 1, 1)$ part of the connection A acts as a vielbein and maps vectors in the tangent space at a point x in spacetime to vectors in the tangent space at $p(x)$ in the internal de Sitter space, given in components with respect to an orthonormal basis at $p(x)$. In order to give their components in the coordinate-induced basis $\{\frac{\partial}{\partial p^a}\}$, we need another vielbein, which can be obtained from (3.4.9) by setting $\omega_0 = e_0 = 0$ (corresponding to spacetime being de Sitter space with cosmological constant Λ) and $p^a(x) = \frac{\kappa}{l}x^a$, as in [153]. We obtain

$$l_n^a(p(x)) = \kappa^2 \frac{\delta_n^a(p \cdot p)\gamma(p) - p^a p_n(\gamma(p) - 1)}{(p \cdot p)(p \cdot p + \kappa^2)}, \quad (3.4.13)$$

where n is a coordinate index in the internal space and a denotes a Lorentz index, as before. This vielbein is of course independent of the underlying spacetime.

Parallel transport can be defined for the $\mathfrak{so}(d, 1)$ connection using the notion of development, which generalises the usual covariant derivative. One introduces a development operator [153]

$$D = d - \frac{1}{2}\omega^{ab}M_{ab} - (e \cdot V), \quad (3.4.14)$$

where the second term is the usual infinitesimal relative rotation of tangent spaces at different spacetime points, and the last term compensates for the change of point of tangency and hence generates maps from the tangent space at one point of de Sitter space to the tangent space at a different point of de Sitter space. Again one should think of an internal space rolled along spacetime [166].

In components, in our conventions we have

$$(\omega^{ab}M_{ab})^c_d = -2\omega^c_d, \quad (3.4.15)$$

and the combination $e^a V_a$ acts on Lorentz indices as an element of $\mathfrak{so}(d - 1, 1)$, representing the map from one tangent space to another in the respective bases. We use

the result obtained by [153] using the techniques of non-linear realisations, namely that when expressed as an $\mathfrak{so}(d-1, 1)$ matrix,

$$l(e \cdot V) = \kappa s(p)^{-1} (e^a X_a) s(p) - s(p)^{-1} [s(p + \delta p) - s(p)] , \quad (3.4.16)$$

where $s(p)$ is defined according to (3.4.8) and δp is determined from the equation

$$[s(p + \delta p)]^a_5 = [(1 + e^b X_b \kappa) s(p)]^a_5 \quad (3.4.17)$$

where only terms linear in e^a are kept in δp . An explicit calculation shows that

$$\delta p^a = \frac{p^a}{\kappa} (\eta_{bc} e^b p^c) + e^a \kappa , \quad (3.4.18)$$

and hence near $p = 0$, we have $\delta p^a = \kappa e^a$, as expected. We find that $(e \cdot V)$ has components

$$(e \cdot V)^b_c = \frac{\kappa (e^b p_c - e_c p^b) (1 - \gamma(p))}{l(p \cdot p)} . \quad (3.4.19)$$

One then has a notion of holonomy, mapping closed loops in spacetime into the internal space by development. In particular, if one develops the field $p(x)$ describing the point of tangency around an infinitesimal closed loop at x_0 , the developed value will in general differ from the original value at x_0 [153]:

$$\Delta p^a(x_0) \propto T_{\mu\nu}{}^i(x_0) l^a_i(p(x_0)) \oint x^\mu dx^\nu , \quad (3.4.20)$$

where $l^a_i(p(x))$ is the inverse of the vielbein (3.4.13) and $T_{\mu\nu}{}^i$ are the components of the torsion tensor $T = de + \omega \wedge e$. This is because by specifying $p^a(x)$, one locally identifies the “internal” de Sitter spaces, thinking of a tangent de Sitter space rolled along the manifold. In a generic situation, where the manifold \mathcal{M} is not identical to the internal space, such an identification is, strictly speaking, only possible at a point. To understand the geometry, one may picture a sphere being rolled along a plane, as in figure 3.2. In general, when the sphere is rolled around and returns to its initial position on the plane, its point of tangency will not be the original one. This is just the situation for the Lorentzian analogues, Minkowski space and de Sitter space, as we will discuss shortly.

The central result we will try to justify in the following is that, starting from Minkowski spacetime, if we assume the internal de Sitter space is rolled along Minkowski space in a non-trivial way, we obtain a connection with torsion. In our interpretation, this is the only way that “coordinates” can act as translations on momentum space, as one normally assumes when associating the Snyder algebra with a noncommutative spacetime.

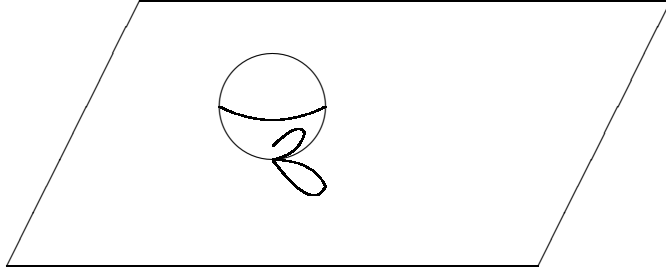


Figure 3.2: When the curved internal space is rolled along Minkowski space, a path in spacetime corresponds to a path in the internal space. Because of the curvature of the internal space, a closed path in Minkowski space does not correspond to a closed path in the internal space, which is manifest as torsion. (We have drawn Riemannian spaces, a plane and a sphere, for clarity of presentation.)

3.5 Cartan Connections on Homogeneous Bundles

This more mathematical introduction into Cartan connections on homogeneous bundles relies mainly on [166], but mentions some additional points which are of importance to our discussion of Einstein-Cartan-Stelle-West theory.

The tangent bundle of a manifold needs to be replaced by a fibre bundle whose fibres are homogeneous spaces $\mathcal{D} \equiv SO(d, 1)/SO(d - 1, 1)$. This can be achieved by starting with a principal bundle $P(\mathcal{M}, SO(d, 1))$, and considering the associated bundle $\mathcal{P} = E(\mathcal{M}, \mathcal{D}, SO(d, 1), P) = P \times_{SO(d, 1)} \mathcal{D}$ (taken as phase space); it can be identified with $P/SO(d - 1, 1)$ by the map

$$\nu : \mathcal{P} \rightarrow P/SO(d - 1, 1), \quad [u, a \cdot SO(d - 1, 1)] \mapsto ua \cdot SO(d - 1, 1). \quad (3.5.1)$$

Then the structure group $SO(d, 1)$ is reducible to $SO(d - 1, 1)$ if the associated bundle \mathcal{P} admits a cross section $\sigma : \mathcal{M} \rightarrow \mathcal{P}$ [99]; furthermore, there is a one-to-one correspondence between reductions of the structure group and cross sections. This cross section, called a *zero section* in [135], corresponds to a choice of origin in the momentum space attached to each point. In physicist's terms, the de Sitter group is spontaneously broken down to the Lorentz group by the choice of points of tangency in the tangent de Sitter spaces at each spacetime point.

The bundle reduction depends on the choice of zero section, or rather, its local representation in coordinates as a function $\mathcal{M} \supset U \rightarrow U \times \mathcal{D}$. This is because the

embedding of $SO(d-1, 1)$ into $SO(d, 1)$ is not canonical, as the stabilisers of different points in \mathcal{D} are isomorphic but related by conjugation. In other words, the mappings appearing in the exact sequence

$$\mathbf{0} \rightarrow SO(d-1, 1) \rightarrow SO(d, 1) \rightarrow SO(d, 1)/SO(d-1, 1) \rightarrow \mathbf{0} \quad (3.5.2)$$

are not canonically chosen (cf. the discussion for the affine group in [135]).

It is of course possible to choose canonical coordinates such that the function representing the zero section is just $x \mapsto (x, [e]) \equiv (x, SO(d-1, 1)) \in \mathcal{M} \times \mathcal{D}$. However, in general we want to locally identify the fibres at nearby base space points, adopting the viewpoint that there is a single tangent \mathcal{D} space which is “rolled along” the manifold. Then we need to retain the general coordinate freedom. (This point is missing in the discussion of [166].) An exact identification is only possible when the connection is flat. Let us assume that coordinates on \mathcal{P} have been fixed, and that it is the zero section, and hence the identification of the fibres, that is varied**. After a choice of zero section, there is still a local gauge freedom corresponding to the stabiliser $SO(d-1, 1)$. We express a given section as $s(x)$, where $\sigma(x) = (x, s(x)) \in \mathcal{M} \times \mathcal{D}$ in our coordinates. The section that corresponds to $s_0(x) \equiv [e]$ will be called “trivial”.

An $\mathfrak{so}(d, 1)$ -valued Ehresmann connection \mathbf{A} in P is in general not reducible to an $\mathfrak{so}(d-1, 1)$ -valued connection in the reduced $SO(d-1, 1)$ bundle $P_R(\mathcal{M}, SO(d-1, 1))$. It can, however, be pulled back using the inclusion

$$\iota_x : SO(d-1, 1) \rightarrow SO(d, 1), \quad \Lambda \mapsto s(x)\Lambda s(x)^{-1} \quad (3.5.3)$$

to a Cartan connection \mathbf{A}_C on the reduced bundle^{††}. Of course reducing the connection to an $\mathfrak{so}(d-1, 1)$ -valued connection and pulling it back to a Cartan connection are very different operations, since in the latter case one wants the $\mathfrak{so}(d, 1)/\mathfrak{so}(d-1, 1)$ part of the pulled-back connection to act as a soldering form, so in particular to be non-singular. We obtain a bundle sequence (cf. [166])

**one is free to choose an “active” or “passive” viewpoint here

††We assume here that the necessary condition $\ker \mathbf{A} \cap (\iota_x)_*(TP_R(\mathcal{M}, SO(d-1, 1))) = \{0\}$ (see [148]) is satisfied.

$$\begin{array}{ccccc}
P_R(\mathcal{M}, SO(d-1, 1)) & \longrightarrow & P(\mathcal{M}, SO(d, 1)) & \longrightarrow & P/\iota_x(SO(d-1, 1))^{\nu_x^{-1}} \simeq \mathcal{P} \\
& \searrow & \downarrow & \swarrow & \\
& & \mathcal{M} & &
\end{array}$$

The reduced bundle $P_R(\mathcal{M}, SO(d-1, 1))$ is mapped into $P(\mathcal{M}, SO(d, 1))$ by

$$\begin{aligned}
p \mapsto [p, e] &= \{(p\Lambda^{-1}, s(x)\Lambda s(x)^{-1}) | \Lambda \in SO(d-1, 1)\} \\
&\in P_R(\mathcal{M}, SO(d-1, 1)) \times_{\iota_x(SO(d-1, 1))} SO(d, 1). \quad (3.5.4)
\end{aligned}$$

The connection one-form A on \mathcal{M} , induced by the connection \mathbf{A} on P , depends on a choice of section $\tau : \mathcal{M} \rightarrow P(\mathcal{M}, SO(d, 1))$. If the zero section σ is fixed, one can choose an arbitrary (local) section $\tau_R : \mathcal{M} \rightarrow P_R(\mathcal{M}, SO(d-1, 1))$ to obtain a section τ ; in local coordinates,

$$\begin{aligned}
\sigma(x) &= (x, s(x)), \quad \tau_R(x) = (x, \Lambda(x)) \\
\longrightarrow \tau(x) &= (x, s(x)\Lambda(x)s(x)^{-1} \cdot s(x)) = (x, s(x)\Lambda(x)). \quad (3.5.5)
\end{aligned}$$

For practical computations, it is often useful to first consider the trivial section. The induced connection corresponding to this section, denoted by $A_0(x)$, is related to the connection for a general section by

$$A(\tau(x)) = \Lambda^{-1}(x)s^{-1}(x)A_0(x)s(x)\Lambda(x) + \Lambda^{-1}(x)s^{-1}(x)ds(x)\Lambda(x) + \Lambda^{-1}(x)d\Lambda(x). \quad (3.5.6)$$

Once the zero section $s(x)$ has been fixed, there is still the freedom of $SO(d-1, 1)$ transformations, corresponding to different choices of $\Lambda(x)$ in (3.5.6). These are the standard local Lorentz transformations in the vielbein formalism of general relativity.

The choice of zero section induces a local splitting of the $\mathfrak{so}(d, 1)$ connection, according to

$$\mathfrak{so}(d, 1) \simeq \mathfrak{so}(d, 1)/\mathfrak{so}(d-1, 1) \oplus \mathfrak{so}(d-1, 1). \quad (3.5.7)$$

This splitting is invariant under the adjoint action of $SO(d-1, 1)$, thus the different parts of the connection will not mix under $SO(d-1, 1)$ transformations. We are considering a reductive geometry (cf. (1.1.6)); recall that the existence of such a splitting

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{h} \quad (3.5.8)$$

invariant under the adjoint action of H , implies that we may identify $\mathfrak{g}/\mathfrak{h}$ with the subspace \mathfrak{x} of \mathfrak{g} . In the present case, this is of course the subspace spanned by the generators $M_{d+1,a}$ of $SO(d, 1)$. For the de Sitter group, just as for the Poincaré and anti de Sitter groups, one can say more: The Lie algebra $\mathfrak{so}(d, 1)$ is equipped with a \mathbb{Z}_2 grading; the subspaces in (3.5.8) satisfy

$$[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{h}, \quad [\mathfrak{h}, \mathfrak{x}] \subset \mathfrak{x}, \quad [\mathfrak{x}, \mathfrak{x}] \subset \mathfrak{h} \quad (3.5.9)$$

and the homogeneous space G/H is a **symmetric space** [167].

The fact that de Sitter space viewed as a homogeneous space is a symmetric space is directly responsible for the splitting of the curvature F of the Cartan connection A into curvature and torsion of the $\mathfrak{so}(d-1, 1)$ connection ω that we have observed above [167]. A general Cartan connection $A = \omega + e$, where ω takes values in \mathfrak{h} and e takes values in \mathfrak{x} , has curvature

$$F = R + \frac{1}{2}[e, e] + d_\omega e, \quad (3.5.10)$$

where $[e, e]$ denotes taking the wedge product on the form part and the commutator on the Lie algebra part, *i.e.*

$$[e, e]_{\mu\nu}^i = C_j^i{}^k (e^j \wedge e^k)_{\mu\nu} = C_j^i{}^k (e_\mu^j e_\nu^k - e_\nu^j e_\mu^k) \quad (3.5.11)$$

if $C_j^i{}^k$ are the structure constants as in (1.1.4). The middle term generically takes values in \mathfrak{h} and \mathfrak{x} , so that one has a splitting

$$F_{\mathfrak{h}} = R + \frac{1}{2}[e, e]_{\mathfrak{h}}, \quad F_{\mathfrak{x}} = \frac{1}{2}[e, e]_{\mathfrak{x}} + d_\omega e. \quad (3.5.12)$$

For a symmetric space, however, $[e, e]_{\mathfrak{x}}$ vanishes [167], and we obtain the result used in the construction of the MacDowell-Mansouri action (3.4.4).

Because we assume A to be a Cartan connection, the $\mathfrak{so}(d, 1)/\mathfrak{so}(d-1, 1)$ part acts as a soldering form, corresponding to the standard vielbein of general relativity; in particular, e_μ^i is an invertible matrix. The soldering form maps vectors in the tangent space $T_x\mathcal{M}$ at a point x in spacetime to vectors in the tangent space $T_{p(x)}\mathcal{D}$ at $p(x)$ in the internal de Sitter space, given in components with respect to an orthonormal basis at $p(x)$. The vielbein that maps between the components of a vector in the orthonormal basis and the coordinate-induced basis is given in (3.4.13).

3.6 Synthesis

The notion of development along curves in spacetime is central to the interpretation of Einstein-Cartan-Stelle-West theory, because it allows “spacetime coordinates” to act as translations in the internal de Sitter space. The situation described by DSR, where noncommuting translations on a curved momentum space are interpreted as noncommuting spacetime coordinates, here corresponds to a Minkowski spacetime with an internal de Sitter space rolled along this Minkowski space. The gauge field $p^a(x)$ specifies the points of tangency of the internal space at each spacetime point, and we have chosen Beltrami coordinates on de Sitter space which look like Cartesian coordinates on Minkowski space near the “origin” of de Sitter space. Since the internal space has a natural scale κ and we needed to introduce a natural scale l in spacetime, we choose the gauge field to be

$$p^a(x) = \frac{\kappa}{l} x^a \quad (3.6.1)$$

in a vicinity of the origin of spacetime which is now taken to be Minkowski space, where x^a are the standard Minkowski coordinates such that the connection vanishes in general relativity. In general a closed path in spacetime will not correspond to a closed path traced out on the internal space, hence such an identification is only local and, strictly speaking, only valid the origin of Minkowski spacetime. On dimensional grounds, the effects of torsion scale as $\frac{x}{l}$ or $\frac{l}{\kappa}$. For (3.6.1) to be well-defined, we must guarantee that $x \cdot x \geq -l^2$, so l should be large in Planck units. We will comment on the significance of the scale l at the end of this section.

It should perhaps be emphasised that the gauge field $p^a(x)$ does not represent physical momentum, but determines the point of tangency of the internal space we have introduced which is to some extent arbitrary. Tangent vectors to the original spacetime can be mapped to tangent vectors to the internal space via the vielbein. The physical interpretation of motion in an internal “momentum” space which is related to motion in spacetime seems obscure, but if coordinates are to act as translations in the internal space, the two must be connected in some way. In this sense, we are constructing the minimal non-trivial gauge field which leads to observable effects, and an alternative interpretation of noncommuting generators X_a in the Snyder algebra.

In our interpretation, different points in the internal de Sitter space do not represent different values for physical four-momentum. Hence we avoid problems with the physical interpretation of DSR, such as the “spectator problem” of noncommutative momentum addition and the “soccer ball problem” of how to describe extended ob-

jects, given that any momentum would appear bounded by the relation $p \cdot p \geq -\kappa^2$. In our framework, tangent vectors representing a particle's (or extended body's) velocity remain vectors and as such live in an unbounded space with commutative addition.

As explained before, we can use equations (3.4.9) to obtain the connection components ω and e corresponding to this choice of our gauge field; we set $\omega_0 = 0$ and $(e_\mu^a)_0 = \delta_\mu^a$ and substitute (3.6.1) to get

$$\begin{aligned}\omega_\mu^{ab} &= (x^a \delta_\mu^b - x^b \delta_\mu^a) \frac{x \cdot x + l^2(\gamma(x) - 1)}{l^2(x \cdot x)\gamma(x)}, \\ e_\mu^i &= \frac{1}{(x \cdot x)(x \cdot x + l^2)} \left(x^i x_\mu (x \cdot x - 2l^2(\gamma(x) - 1)) + \delta_\mu^i (x \cdot x) 2l^2\gamma(x) \right)\end{aligned}\quad (3.6.2)$$

and

$$\begin{aligned}\partial_\nu e_\mu^i - \partial_\mu e_\nu^i &= (x_\nu \delta_\mu^i - x_\mu \delta_\nu^i) \frac{l^2(2l^2(\gamma(x) - 1) - 3(x \cdot x)) - (x \cdot x)^2}{(x \cdot x)(x \cdot x + l^2)^2}, \\ \omega_\nu^{ib} e_{\mu b} - \omega_\mu^{ib} e_{\nu b} &= \frac{(2l^2 + x \cdot x)(x \cdot x + l^2(\gamma(x) - 1))}{(x \cdot x)(x \cdot x + l^2)l^2\gamma(x)} (x_\nu \delta_\mu^i - x_\mu \delta_\nu^i),\end{aligned}\quad (3.6.3)$$

where now

$$\gamma(x) \equiv \sqrt{\frac{x \cdot x + l^2}{l^2}}, \quad (3.6.4)$$

which gives a non-zero torsion

$$T_{\mu\nu}^i = (x_\nu \delta_\mu^i - x_\mu \delta_\nu^i) \frac{1}{l^2 \sqrt{\frac{x \cdot x}{l^2} + 1}}. \quad (3.6.5)$$

Interestingly enough, for the choice of zero section (3.6.1) the scale κ drops out of all expressions. Expressed in coordinates on the internal space, one has

$$T_{\mu\nu}^i = (p_\nu \delta_\mu^i - p_\mu \delta_\nu^i) \frac{1}{l\kappa \sqrt{\frac{p \cdot p}{\kappa^2} + 1}}. \quad (3.6.6)$$

The quantity $T_{\mu\nu}^i$ will be multiplied by an infinitesimal closed loop $\oint x^\mu dx^\nu$ to give the difference in the value $p(x)$ caused by development along this loop. In momentum coordinates, this is equal to $\frac{l}{\kappa} \oint p^\mu dp^\nu$, and the effect of going around the developed curve in the internal space is (near $x = 0$ or $p = 0$) proportional to κ^{-2} , just as was suggested by (3.2.2).

Expressing Minkowski space in the usual coordinates, together with the (local) identification $p^a(x) = \frac{\kappa}{l} x^a$, in this framework gives a connection with torsion. Developing a closed curve in spacetime in the internal space will give a curve that does not close in general, which is the effect of noncommuting translations in the internal space.

The reader may wonder how the “deformation” of the Minkowski solution described here is manifest in a metric. We can define a metric by the usual expression

$$g_{\mu\nu} = e_\mu^a e_\nu^b \eta_{ab}. \quad (3.6.7)$$

This metric would not determine the connection, but could be used to define distances in the spacetime in the usual way. Then, from (3.6.2), we get

$$g_{\mu\nu} = \eta_{\mu\nu} \frac{4}{1 + \frac{x \cdot x}{l^2}} + x_\mu x_\nu \frac{(x \cdot x)}{((x \cdot x) + l^2)^2}. \quad (3.6.8)$$

It should be stressed that the connection on spacetime is *not* the Levi-Civita connection of this metric, as we are working in a first order formulation where metric and connection are independent. There is a factor of 4 because of a term in (3.4.12) which does not necessarily go to zero as $p \rightarrow 0$. With the identification (3.6.1), the soldering form always gets a contribution

$$e_\mu^i(x) = (e_\mu^i)_0 + \delta_\mu^i + O\left(\frac{x}{l}\right). \quad (3.6.9)$$

The limit $\kappa \rightarrow \infty$ is now identified with the limit $l \rightarrow \infty$, in which we recover the (rescaled) Minkowski metric.

In deriving the expressions (3.6.2) we started with Minkowski space, which clearly solves the field equations of the Einstein-Cartan theory for an energy-momentum tensor cancelling the cosmological constant term, and vanishing internal spin. In changing the zero section, we then performed a $SO(d, 1)$ gauge transformation, under which the curvature F transformed as

$$F(s(x)) = s^{-1}(x)F(x)s(x). \quad (3.6.10)$$

This is a general $SO(d, 1)$ rotation which mixes up the $\mathfrak{so}(d-1, 1)$ and $\mathfrak{so}(d, 1)/\mathfrak{so}(d-1, 1)$ parts of the connection and the curvature. Hence, the resulting connection will no longer solve the original field equations, but the field equations for an energy-momentum tensor which has also undergone a $SO(d, 1)$ transformation. This mixes up the energy-momentum and internal spin parts, combining them into an element of the Lie algebra $\mathfrak{so}(d, 1)$, the interpretation of which seems obscure at least.

A comment is in order with regard to physical units. In addition to the energy scale κ , which is perhaps naturally identified with the Planck scale, the identification of lengths with momenta, necessary in the framework presented here, requires the choice of a unit of length l which is not necessarily connected to the scale κ . It may

well be that it is instead the cosmological constant which sets this length scale, leading to an astronomical scale instead of a sub-atomic one. And indeed, some more recent approaches to quantum gravity (*e.g.* [58]) use the product $G\Lambda$ as a dimensionless parameter in a perturbative expansion. A fixed positive Λ also seems to be required in non-perturbative approaches to quantum gravity [6]. Then the cosmological constant may play the role of a fundamental parameter in quantum gravity.

3.7 Summary and Discussion

It has been argued that the algebra of DSR describes the symmetries of a semiclassical limit of (a generic theory of) quantum gravity (see *e.g.* [6]). If this claim is taken seriously, one has to give an interpretation of the noncommuting translations appearing in the algebra, and usually they are supposed to represent a spacetime with a fundamentally noncommutative structure [111]. Alternatively, one may view the apparent noncommutativity as an artefact of the finite resolution of lengths [70]. However, there are fundamental difficulties in associating these noncommuting operators directly with coordinates on spacetime, as position is not *additive* in a way that momentum and angular momentum are [37]. Furthermore, as also pointed out in [37], a proposed noncommutativity of spacetime of the form (3.2.2), proportional to angular momentum or boost generators, and hence vanishing at a given “origin”, seems deeply at odds with any idea of (even Galilean) relativity. This would also be an obvious criticism of the framework presented in this chapter, when taken as a theory that is supposed to describe the real world.

What we have shown here, is that using the framework of Einstein-Cartan-Stelle-West theory, one reaches a different conclusion from the usual one: The noncommutativity of translations on a momentum space of constant curvature is interpreted as torsion of a connection that solves the equations of Einstein-Cartan theory with a modified energy-momentum tensor that mixes with the spin tensor. If one takes this seriously, one is led to conclude that there is an effect of torsion induced by quantum gravity, whose effects would however only become measurable over distances comparable to l , a length scale presumably associated with the cosmological constant.

No such effect appears in de Sitter space with an appropriate cosmological constant, or indeed any vacuum solution of the theory. Vacuum solutions are then just described by the Poincaré algebra, and hence undeformed special relativity. It is curi-

ous that the curvature in the internal “momentum” space is reflected in torsion when Minkowski spacetime is considered as spacetime, which would however seem rather unnatural as it is not a vacuum solution any more. This result does however seem to be the most natural interpretation of the DSR algebra in a differential-geometric framework capable of describing more general spacetimes.

Any non-zero energy-momentum tensor, however, will lead to a connection having torsion. In theories such as Einstein-Cartan theory, this leads to well-known problems when trying to couple the gravitational field to Maxwell fields, for instance, as there is no unambiguous procedure of minimal coupling. This is because the statement that the exterior derivative is independent of the choice of connection,

$$(dA)_{\mu\nu} \propto \partial_{[\mu} A_{\nu]} = \nabla_{[\mu} A_{\nu]}, \quad (3.7.1)$$

is true precisely when torsion vanishes. Using an $\mathfrak{so}(d-1, 1)$ connection, this is apparent from

$$d(e^i A_i) = \nabla(e^i A_i) = A_i \nabla e^i - e^i \wedge \nabla A_i = -e^i \wedge \nabla A_i + A_i T^i \quad (3.7.2)$$

where $\nabla e^i = de^i + \omega^i_j \wedge e^j$ etc. One has two different candidates for the field strength F , namely $e^i \wedge \nabla A_i$ and $d(e^i A_i)$, with possibly observable differences between these choices, although it could be argued that $F = dA$ is the only meaningful choice because it preserves gauge invariance [21].

In the framework of Einstein-Cartan-Stelle-West theory, gauge fields should be coupled to gravity via development, *i.e.* replacing $F = dA$ by $F = DA$. We compute from (3.4.9) and (3.4.19) that development can be expressed in terms of ω_0 and e_0 by

$$D = d + \omega - (e \cdot V) = d + \omega_0 + 2(p \otimes_A e_0) \kappa \frac{(\gamma(p) - 1)}{l(p \cdot p)} =: d + \omega_{\text{eff}}, \quad (3.7.3)$$

where \otimes_A is an antisymmetrised tensor product, $(U \otimes_A V)^{ab} \equiv U^{[a} V^{b]} \equiv \frac{1}{2}(U^a V^b - U^b V^a)$. Parallel transport is effectively described by the connection ω_{eff} , whose torsion is in general non-zero. One can give an explicit formula for the torsion which is however rather complicated and does not seem to give much insight; to linear order in p^i , one has

$$T^i = \frac{l}{\kappa} (R^i_b)_0 p^b - \frac{1}{2l\kappa} e_0^i \wedge (e_j^0 p^j) + \frac{1}{2\kappa^2} (p^i e_{j0} \wedge dp^j - p_j e_0^i \wedge dp^j) + O(p^2). \quad (3.7.4)$$

If we assume a universal relation of internal momenta and spacetime lengths of the form $p \sim \frac{\kappa}{l} x$, the second and third terms seem to give contributions of order x/l^2 . The

first term is proportional to the local curvature of ω_0 , $R^i_b = d\omega^i_b + \omega^i_j \wedge \omega^j_b$, contracted with x^b . Note that it is the Riemann tensor, not the Ricci tensor, that appears, so that propagating degrees of freedom of the gravitational field are included. This first term should in realistic situations, even in vacuum, give the dominant contribution.

Assuming that minimal coupling is achieved through the development operator D , or equivalently by using the effective connection which has torsion, one would couple vector or matter fields (using $D\psi$ for spinors) to torsion, breaking gauge invariance. Such an effect of course leads to the absence of charge conservation, and this should be experimentally observable in the presence of a non-trivial gravitational field, *i.e.* in regions where spacetime is not exactly de Sitter. Let us recall that in standard tensor calculus one uses the identity

$$[\nabla_\mu, \nabla_\nu]M_{\lambda\rho} = R_{\mu\nu\lambda}{}^\sigma M_{\sigma\rho} + R_{\mu\nu\rho}{}^\sigma M_{\sigma\lambda} - T_\mu{}^\sigma{}_\nu \nabla_\sigma M_{\lambda\rho} \quad (3.7.5)$$

which gives for an antisymmetric $M_{\lambda\rho}$ when contracted

$$[\nabla^\lambda, \nabla^\rho]M_{\lambda\rho} = -g^{\mu\lambda}g^{\nu\rho}T_\mu{}^\sigma{}_\nu \nabla_\sigma M_{\lambda\rho}, \quad (3.7.6)$$

to establish that the right-hand side of Maxwell's equation $\nabla^\lambda F_{\lambda\rho} = 4\pi J_\rho$ satisfies a continuity equation in the absence of torsion. With torsion present, one has then for any region R

$$\int_{\partial R} d^3x \sqrt{h} n^\lambda J_\lambda = \frac{1}{4\pi} \int_R d^4x \sqrt{g} (-g^{\mu\lambda}g^{\nu\rho}T_\mu{}^\sigma{}_\nu \nabla_\sigma F_{\lambda\rho}). \quad (3.7.7)$$

Effects become important when the size of the region R is comparable to the length scale of torsion.

As an example consider the Schwarzschild solution, which has Kretschmann scalar

$$R_{abcd}R^{abcd} \sim \frac{r_S^2}{r^6}, \quad (3.7.8)$$

so roughly $R_{abcd} \sim r_S r^{-3}$. Assuming that the origin of the x coordinate system corresponds to the centre of the Earth, we would, on the surface of the Earth, measure a torsion of order $r_S R^{-2}$, where R is the radius of the Earth. With $r_S \sim 10^{-2}$ m and $R^2 \sim 10^{13}$ m², the length scale for effects of torsion would be about 10^{11} m. The other two contributions, given that $l \sim 10^{26}$ m, would be much smaller. Although this crude estimate suggests that effects will be very small, even tiny violations of charge conservation should have been observed experimentally. For a discussion of experimental tests of charge conservation and possible extensions of Maxwell theory

in Minkowski space, see [102]. Processes such as electron decay on a length scale of 10^{11} m, or a time scale of 10^3 s, can clearly be ruled out.

The example presented here shows that the correct physical interpretation of purely algebraic relations, such as the commutators of the Snyder algebra, may not be the seemingly obvious one. We conclude that the physical motivation for assuming spacetime is “noncommutative” may not be as clear as often assumed.

3.8 Gauge Invariance Broken?

When discussing the issue of possible consequences of broken gauge invariance, we must bear several points in mind.

The idea that an asymmetry between the proton and electron charges could have interesting astrophysical consequences goes back to Lyttleton and Bondi [108], who argued that a charge difference, and hence a net charge of the hydrogen atom, of 10^{-18} elementary charges, might explain the observed expansion of the universe by electrostatic repulsion. This idea was proposed in connection with Hoyle’s ideas of a universe in a steady state, which required continuous production of material via a “creation field” [85], and a modification of Maxwell’s equations was proposed to accommodate charge nonconservation. From Hoyle’s perspective, however, the steady state model was incompatible with expansion of the universe by electrostatic repulsion, and should lead to electrostatic attraction instead [86].

There seems to be no need for the electron and proton charges to be of equal magnitude to maintain gauge invariance. However, if the universe as a whole is not neutral, but it is homogeneous, gauge invariance must be broken. Hence the two issues are closely related. Modern laboratory experiments [115] give a bound of 10^{-21} elementary charges on the difference of electron and proton charge; astrophysical considerations give bounds of 10^{-26} elementary charges using the isotropy of the cosmic microwave background [30], or 10^{-29} elementary charges by considering cosmic rays [124]. Recently, an interesting proposal to measure net charges of atoms and neutrons, sensitive to 10^{-28} elementary charges, was put forward [8].

From a theoretical viewpoint, if gauge invariance is broken, it is natural to assume a nonvanishing photon mass. One then considers Einstein-Proca theory, an outline of which can be found in [79]. The photon may also be charged. Here, experimental bounds on the charge are 10^{-29} elementary charges using pulsars [141], and possibly

10^{-35} elementary charges from CMB isotropy [30].

Experimental bounds on violations of gauge invariance in electrodynamics are very tight, and hence any theory predicting torsion which is coupled to electromagnetism faces severe problems when confronted by experiment. In the framework of Einstein-Cartan-Stelle-West theory, it is possible to maintain gauge invariance by choosing $F = dA$, but using the development operator is the most natural choice.

3.9 Symmetric Affine Theory

If Einstein-Cartan theory is considered as the extension of general relativity which allows for torsion, there is an analogous extension which allows for a non-metric connection. This theory can be formulated in terms of a torsion-free $\mathfrak{gl}(n, \mathbb{R})$ connection and is known as **symmetric affine theory**. It is equivalent to standard general relativity with a massive vector field, known as (nonlinear) Einstein-Proca theory [79].

One could attempt to embed this theory into a theory of a connection taking values in the algebra of the affine group $\mathfrak{a}(n, \mathbb{R})^{\sharp\sharp}$,

$$A = \begin{pmatrix} \omega^a_b & \frac{1}{l}e^i \\ 0 & 0 \end{pmatrix}, \quad (3.9.1)$$

where now ω^a_b is not constrained by $\omega^{ab} = -\omega^{ba}$. Geometrically, this means that the connection does not preserve the lengths of vectors under parallel transport.

The corresponding curvature of A (R is the curvature of the $\mathfrak{gl}(n, \mathbb{R})$ part),

$$F = dA + A \wedge A = \begin{pmatrix} R^a_b & \frac{1}{l}T^i \equiv \frac{1}{l}(de^i + \omega^i_j \wedge e^j) \\ 0 & 0 \end{pmatrix}, \quad (3.9.2)$$

would then be constrained by demanding that $T^i \equiv 0$. This seems rather unnatural from the perspective of Cartan geometry. Furthermore, the length scale l is now completely arbitrary as it does not appear in the $\mathfrak{gl}(n, \mathbb{R})$ part of the curvature any more, just as for a connection taking values in the Poincaré algebra.

^{$\sharp\sharp$} For a comprehensive review of general theories of this type, see [78].

One proceeds by considering Lagrangians that only depend on the Ricci tensor, which is a one-form obtained by contracting the components of the Riemann curvature, written in the basis of one-forms given by the vielbein e^i :

$$\text{Ric}_a = \text{Ric}_{ia} e^i, \quad \text{Ric}_{ia} = R_{ji}{}^j{}_a, \quad (3.9.3)$$

where the curvature two-form is

$$R^a{}_b = \frac{1}{2} R_{ij}{}^a{}_b e^i \wedge e^j. \quad (3.9.4)$$

One then splits the Ricci tensor into symmetric and antisymmetric part, symmetrising over a component (with respect to the given basis) index and a $\mathfrak{gl}(n, \mathbb{R})$ index. The antisymmetric part can be interpreted as a spacetime two-form

$$i_{e_a} (R^a{}_b \wedge e^b), \quad (3.9.5)$$

where i_{e_a} is interior multiplication with the vector e_a , defined by being dual to the one-forms e^b :

$$e^b(e_a) = \delta^b_a. \quad (3.9.6)$$

No such construction is possible for the symmetric part, which is normally more relevant in concrete constructions. The splitting itself seems depend on the choice of basis.

Chapter 4

General Relativity From a Constrained Topological Theory

“What, then, is time? If no one ask me, I know; if I wish to explain to him who asks, I know not.” [11]

St. Augustine

4.1 Overview

In this chapter, we shall investigate another possible formulation of general relativity in four dimensions, in terms of a constrained topological theory. Here the term “topological” means that the theory has a very large symmetry: All of its solutions are locally gauge equivalent, and there are no local degrees of freedom. Globally, not all solutions are necessarily gauge equivalent if there are topological obstructions. Starting from such a theory, we will use constraints that restrict the possible gauge transformations and hence allow for a less trivial theory, namely one with local degrees of freedom that are those of general relativity.

The topological theory we consider is usually known as BF theory, deriving its name from the action

$$S = \int B^{ab} \wedge F_{ab}[A], \quad (4.1.1)$$

where F is the curvature of a G -connection A and B is a \mathfrak{g} -valued two-form (in our case G will be a rotation group). The equations of motion are

$$F^{ab} = 0, \quad \nabla B^{ab} = 0, \quad (4.1.2)$$

where ∇ is a covariant exterior derivative involving the connection A . All flat connections are locally gauge equivalent, and the (local) equivalence of all solutions to the second equation can be seen [13] by noting that the transformation

$$B^{ab} \mapsto B^{ab} + \nabla \eta^{ab}, \quad (4.1.3)$$

where η^{ab} is a \mathfrak{g} -valued one-form, is a symmetry of the action (4.1.1). This transformation is generated by the equation $F^{ab} = 0$, in a canonical sense, and should be viewed as a generalised gauge transformation. It is a symmetry on the phase space of the theory. Locally, any solution to $\nabla B^{ab} = 0$ is of the form $\nabla \eta^{ab}$, so any solution is gauge equivalent to $B^{ab} = 0$ in this generalised sense.

The motivation for considering BF theory is that general relativity in three dimensions is just of this form: The first order formulation can be written as

$$S = \int \epsilon_{abc} (e^a \wedge R^{bc}[\omega]) , \quad (4.1.4)$$

and hence general relativity in three dimensions is topological. Since the quantisation of this system is relatively well understood (see [168] for Witten's formulation as a Chern-Simons theory, and the book [31] for an overview over different routes to quantisation), and seems to have many conceptual aspects such as diffeomorphism and Lorentz invariance in common with the higher-dimensional case, one might hope to gain understanding of higher-dimensional general relativity from studying BF theory. On the other hand, there is of course a fundamental difference between a system with an infinite number of (local) degrees of freedom and one with only (usually a finite number of) global degrees of freedom, which makes it rather unclear whether this approach will be fruitful.

It has been shown [56] that in any number of dimensions $d \geq 4$ general relativity can be formulated as a constrained BF theory, where one adds to the BF action constraints quadratic in the field B^{ab} , restricting it to be a wedge product of one-forms e^a that can be interpreted as a frame field. In four dimensions, where one has to enforce

$$B^{ab} = \epsilon^{ab}_{cd} e^c \wedge e^d \quad (4.1.5)$$

to reproduce the first order GR action (3.1.1), this formulation (in slightly different complex form) was first studied by Plebanski [138]. This is also the situation most studied in the spin foam approach to quantum gravity, where one tries to quantise general relativity by quantising a BF-type theory and then implementing constraints.

In this chapter, we shall investigate a formulation which is in the spirit of the Plebanski formulation, but involving only linear constraints, of the type used recently as a new idea in the spin foam approach to quantum gravity. We identify both the continuum version of the linear simplicity constraints used in the quantum discrete context and a linear version of the quadratic volume constraints that are necessary to complete the reduction from the topological theory to gravity. We illustrate and discuss also the discrete counterpart of the same continuum linear constraints. Moreover, we show under which additional conditions the discrete volume constraints follow from the simplicity constraints, thus playing the role of secondary constraints.

In the general context of the thesis, we see how a very large symmetry present in the original action (4.1.1) can be constrained appropriately so as to give a theory with rather different physics – one with local degrees of freedom. This symmetry acts neither on spacetime nor on fibre bundles, but on a symplectic manifold, namely the phase space of a dynamical system.

4.2 Actions for Gravity

It should not be surprising that the equations of general relativity can be derived from several different action principles, leading to equivalent classical theories (in the case of pure gravity, at least). The statement that given equations of motion arise from different actions is the functional equivalent of saying that there are many functions with the same stationary points. Apart from constructing different actions as functionals of the same variables, one can also express the same physical content in terms of different variables (by performing a Legendre transformation, or otherwise), leading to even more possibilities for putting the same physics into a different mathematical form.

This introduces several possible sources of ambiguity into quantisation; Feynman's path integral approach to quantum mechanics is based on the fact that off-shell configurations, *i.e.* those that do not solve the equations of motion and are not stationary points of the action, contribute to transition amplitudes as well. Hence different functionals of the same variables, while giving the same equations of motion, may give different quantum theories. For formulations in terms of different variables it is even less clear whether the resulting quantum theories are related or not. One may see this as a problem or as a virtue; when a classical theory such as general relativity seems

to resist conventional quantisation techniques, one may hope that one of the possible classical reformulations offers more promise for quantisation. The most obvious example of this is Ashtekar new variables [10] as a possible new route to quantising general relativity in four dimensions. In three dimensions, a number of consistent quantisation schemes for general relativity are known, leading to (at least apparently) inequivalent theories [31].

But already classically, in the case of general relativity equivalent formulations for pure gravity can give non-equivalent theories when coupled to matter. We encountered an instance of this in chapter 3 when discussing Einstein-Cartan theory; in contrast to conventional general relativity, where torsion is set to zero by *flat*, in Einstein-Cartan theory matter can act as a source of torsion. This is a non-equivalence on the level of equations of motion which can therefore be tested more easily by experiment. In this chapter we shall consider a reformulation of pure general relativity only.

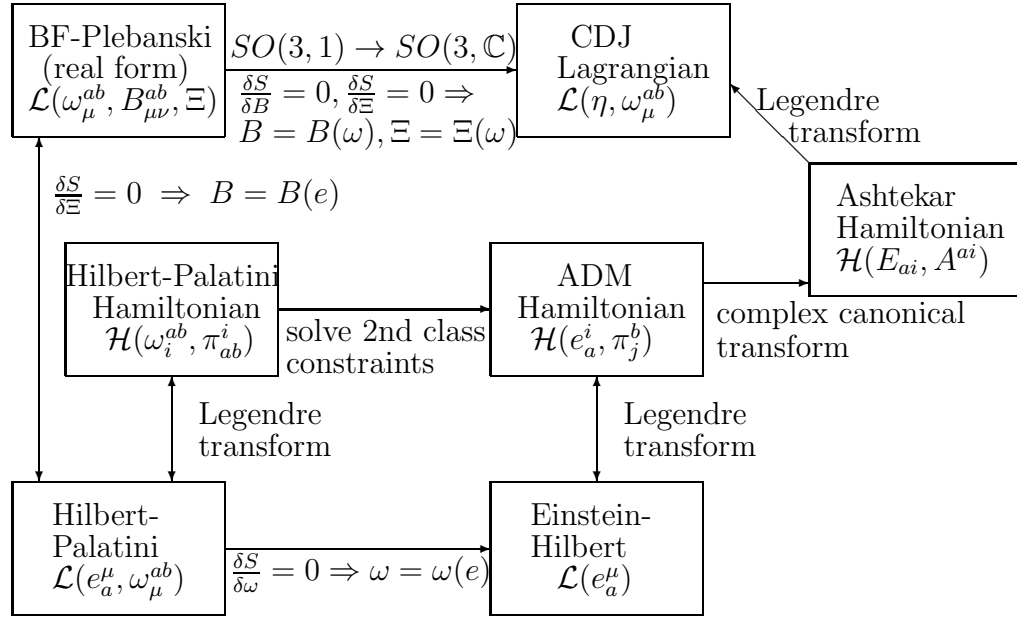


Figure 4.1: Lagrangians and Hamiltonians for general relativity in four dimensions. (Essentially taken from [131].)

In figure 4.1, essentially taken from the review [131], we show a few possible action principles for general relativity and their interrelations.* The most commonly used

*Figure 4.1 is by no means meant to be a complete display of all known actions for general

ones are the Einstein-Hilbert action [82], the Palatini first order formulation, and its modification proposed by Holst [83]. This last one is of special interest because it is the classical, covariant starting point for the canonical quantisation leading to loop quantum gravity [145]. Even if not the only possible useful one [150], a particularly popular action in covariant approaches to quantising gravity [122], like the spin foam [133] and group field theory approach [121], is the formulation as a constrained BF (or Plebanski [138, 101]) theory. Here one starts from topological BF theory [84] in four spacetime dimensions, and adds suitable constraints on the two-form B variables of the theory such that, on solutions of these constraints, the action reduces to the Palatini or Holst action for general relativity. We will summarise the idea behind this formulation in the following. In the original Plebanski formulation the constraints on the B variables are quadratic, and so are the discrete constraints that are then implemented in the spin foam models based on a simplicial discretisation. On the other hand, the most recent developments in the spin foam and group field theory approach to quantum gravity are based on a linear set of discrete constraints, which can be shown to be slightly stronger, in the restrictions they impose on the original BF configurations, than some of the original discrete quadratic constraints. Once more, we will detail this construction in the following.

Here we investigate whether a formulation in terms of linear constraints is also possible in the classical continuum theory, and what it implies. We will see that the replacement of the so-called diagonal and cross-simplicity constraints with linear constraints at the continuum level is relatively straightforward, after one has introduced new variables n_a forming a basis of three-forms at each point. One then needs additional constraints corresponding to the volume constraints. We will see that one can linearise these constraints too. We then give a discrete version of these linear volume constraints (thinking of spacetime which is discretised in a triangulation), which bears a striking resemblance to the so-called “edge simplicity” constraints of [47]. We note that only certain linear combinations of the volume constraints one would naively write down are necessary to constrain the bivectors $\Sigma^{AB}(\triangle)$ sufficiently. Similarly to the quadratic case, we will also see that when linear diagonal and off-diagonal constraints hold everywhere in a 4-simplex, and one also imposes “closure” constraints on both bivectors $\Sigma^{ab}(\triangle)$ (associated to triangles) and normals $n_a(\triangle)$ (associated to relativity; for instance, an interesting variant of the (complex) Plebanski formulation has recently been explored by Eyo Eyo Ita III (Ph.D. thesis, University of Cambridge (2010))).

tetrahedra), the sufficient set of linear combinations of the linear volume constraints follows. This additional four-dimensional closure constraint on the normal vectors has, to the best of our knowledge, not been considered or implemented as an additional condition in the spin foam literature yet, although it does appear in some first order formulation of Regge calculus [35], and it plays also a role in the discrete analysis of [40].

4.3 Constraining BF Theory

Let us briefly review what is known at the classical continuum and discrete level, concerning the Plebanski formulation of classical gravity. We limit our considerations to the covariant, Lagrangian context, and to a very small subset of the available results, those which have been already of direct relevance for quantum gravity model building, especially in the spin foam context. For recent results in the canonical Hamiltonian setting, both continuum and discrete, see [28, 47, 160].

In this chapter we consider the Einstein-Hilbert-Palatini-Holst [82, 83] Lagrangian (without cosmological constant), a modification of the action (3.1.1) used earlier,

$$S_{\text{EHPH}} = \frac{1}{8\pi G} \int_{\mathcal{R} \times \mathbb{R}} \left(\frac{1}{2} \epsilon_{abcd} e^a \wedge e^b \wedge R^{cd}[\omega] + \frac{1}{\gamma} e^a \wedge e^b \wedge R_{ab}[\omega] \right), \quad (4.3.1)$$

where spacetime is assumed to be of the form $\mathcal{R} \times \mathbb{R}$ so that a (3+1) splitting can be performed, ω^{ab} is a G -connection one-form (the gauge group G is $SO(3,1)$ or $SO(4)$, or an appropriate cover), R^{ab} its curvature, and e^a is an \mathbb{R}^4 -valued one-form representing an orthonormal frame. The term involving γ , known as the Holst term, is not relevant classically; it is, up to a total derivative[†], proportional to $T^a \wedge T_a$, and hence does not modify the classical equation of motion $T^a = 0$. When (4.3.1) is coupled to matter, the Holst term leads to a rescaling of the coupling of the matter spin density to torsion.

Though not relevant classically, the Holst term is of fundamental importance in loop quantum gravity (LQG), where γ is known as the Barbero-Immirzi parameter, and more generally for any canonical formulation of gravity, as it modifies the symplectic structure of the theory.

If one introduces a \mathfrak{g} -valued two-form,

$$B^{ab} = \frac{1}{8\pi G} \left(\frac{1}{2} \epsilon^{ab}_{cd} e^c \wedge e^d + \frac{1}{\gamma} e^a \wedge e^b \right), \quad (4.3.2)$$

[†]For a discussion of the total derivative term see [117].

then the action (4.3.1) becomes[‡]

$$S = \int B^{ab} \wedge R_{ab}[\omega] + \lambda^\alpha C_\alpha[B], \quad (4.3.3)$$

i.e. it takes the form of a topological BF theory with additional constraints C_α which enforce that B^{ab} is indeed of the form (4.3.2), and that are enforced by means of Lagrange multipliers λ_α .

As said, BF theory without constraints is topological. Its equations of motion imply that ω^{ab} is flat and the covariant exterior derivative of B^{ab} vanishes. Having no local degrees of freedom, the quantisation of such a theory is therefore rather simple and quite well understood. Inspired by this classical formulation, the main issue when trying to construct a quantum theory related to quantum gravity, in four dimensions, is then the correct implementation of appropriate constraints that lead to (4.3.2) for some set of one-forms e^a , either at the level of quantum states or in a path integral formulation. Indeed, the bulk of the work in the spin foam approach [49, 57, 105, 133] (as well as in the group field theory formalism [16, 121, 123]), in recent years, has been devoted to this task. These constraints are also the subject of this chapter.

To simplify the following calculations, we introduce another two-form field Σ^{ab} ,

$$\Sigma^{ab} \equiv \frac{1}{1 - s\gamma^2} \left(B^{ab} - \frac{\gamma}{2} \epsilon^{ab}_{cd} B^{cd} \right), \quad (4.3.4)$$

where s is the spacetime signature, $s = -1$ for $G = SO(3,1)$ and $s = +1$ for $G = SO(4)$ (and we assume $\gamma^2 \neq s$)[§]. This is a linear redefinition which simplifies the constraint (4.3.2),

$$\Sigma^{ab} = \frac{1}{8\pi\gamma G} e^a \wedge e^b, \quad (4.3.5)$$

but leads to more terms in the action. The translation of all calculations from one set of variables to the other is usually straightforward.

The traditional way to enforce the restriction (4.3.5), the one matching the original classical Plebanski formulation of gravity, was to add quadratic **simplicity constraints**[¶] to the action [45, 143],

$$\epsilon_{abcd} \Sigma^{ab}_{\mu\nu} \Sigma^{cd}_{\rho\sigma} = V \epsilon_{\mu\nu\rho\sigma}, \quad (4.3.6)$$

[‡]We use indices from the beginning of the Greek alphabet such as α to denote abstract indices not associated with certain transformation groups.

[§]For uniformity of the discussion, we shall in the following talk about “time” and use the label 0 even when the gauge group is $SO(4)$ and the signature Riemannian.

[¶]“Simplicity” because a two-form that can be written as a wedge product of one-forms is called simple.

where V can be expressed in terms of Σ^{ab} by contracting (4.3.6) with $\epsilon^{\mu\nu\rho\sigma}$, to give: $V = \frac{s}{24}\epsilon^{\mu\nu\rho\sigma}\epsilon_{abcd}\Sigma_{\mu\nu}^{ab}\Sigma_{\rho\sigma}^{cd}$. This is itself a reformulation of the original Plebanski constraint, which would read:

$$\epsilon^{\mu\nu\rho\sigma}\Sigma_{\mu\nu}^{ab}\Sigma_{\rho\sigma}^{cd} = V\epsilon^{abcd}, \quad (4.3.7)$$

and is equivalent to the first under assumption that $V \neq 0$ everywhere. The version (4.3.6) has the advantage of permitting a much simpler discretisation and thus a more straightforward implementation within the spin foam formalism. Under the same assumption $V \neq 0$, there are the following four classes of solutions to (4.3.6):

$$\text{either } \Sigma^{ab} = \pm e^a \wedge e^b \quad \text{or} \quad \Sigma^{ab} = \pm \frac{1}{2}\epsilon^{ab}_{cd}E^c \wedge E^d \quad (4.3.8)$$

for some set of one-forms E^a or e^a (in the following, we reserve e^a for one-forms satisfying the first relation, viewed as encoding the metric; since we consider pure gravity, the factor $8\pi\gamma G$ can obviously be introduced by rescaling). One would like to select only the first class of solutions $\Sigma^{ab} = +e^a \wedge e^b$, which, when substituted in the BF action, gives the Holst action (4.3.1). Classically, this is not a severe problem. As shown in [143], non-degenerate initial data of a solution of the form $\Sigma^{ab} = +e^a \wedge e^b$ generically remain within the same branch of solutions. The situation in the quantum theory, where one necessarily has contributions from all branches, is less clear.

More troublesome, if $V = 0$, the field Σ^{ab} does not permit a straightforward geometric interpretation at all. Since in the region of the phase space where $V = 0$, the theory is less constrained, and hence has more degrees of freedom, these non-geometric configurations should be expected to be dominating in a path integral [143], unless measure factors are such that this is avoided.

Spin foam models are usually defined in a piecewise flat context, and spin foam amplitudes are defined for given simplicial complexes [133]. Therefore one is interested in identifying a discrete version of the above constraints that could be imposed at the level of each complex. The version (4.3.7) of the simplicity constraints admits only a rather involved discrete counterpart [45] and, upon quantisation, leads to the Reisenberger model [45, 142], which has so far received only limited attention.

The discrete analogue of the constraints (4.3.6) led instead [45, 143] to the construction of the Barrett-Crane model [17], in the case in which the Barbero-Immirzi parameter is excluded from the original action ($\gamma \rightarrow \infty$). The construction is initially limited to a single 4-simplex, the convex hull of five points in \mathbb{R}^4 ($\mathbb{R}^{1,3}$, in the Lorentzian case) with the topology of a 4-ball, whose boundary is triangulated by the

five tetrahedra identified by the five independent subsets of four such points, while subsets of three points identify the four triangles belonging to each of these five tetrahedra, each of the triangles being shared by a pair of tetrahedra (see figure 4.2). One

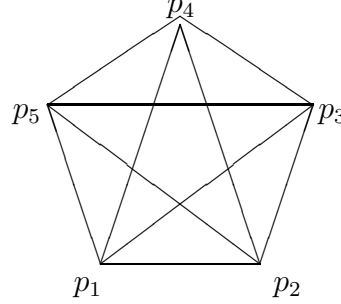


Figure 4.2: A 4-simplex is generated by five points.

then associates a Lie algebra element (bivector) $\Sigma_{\Delta}^{ab} \in \mathfrak{so}(4) \simeq \wedge^2 \mathbb{R}^4$ (similarly in the Lorentzian case) to each triangle Δ in a given triangulation by integrating the two-form Σ^{ab} over Δ . The task is then to constrain appropriately these Lie algebra variables (or their quantum counterpart) following the continuum treatment.

It is useful to split the set of continuum equations (4.3.6) into two sets. Out of the 21 equations (4.3.6), one first identifies and imposes those 18 which have zero on the right-hand side (the “diagonal” and “off-diagonal” simplicity constraints),

$$\epsilon_{abcd} \Sigma_{\mu\nu}^{ab} \Sigma_{\mu\nu}^{cd} = \epsilon_{abcd} \Sigma_{\mu\nu}^{ab} \Sigma_{\mu\rho}^{cd} = 0 \quad \forall \mu, \nu, \rho \text{ (no sum over } \mu, \nu). \quad (4.3.9)$$

This corresponds to the case if one or two of the indices of the two fields Σ coincide. At the discrete level, this translates into two triangles on which the same fields are discretised which either coincide or at least share a single edge, and thus belong to the same tetrahedron. Thus all bivectors Σ_{Δ}^{ab} are required to satisfy

$$\epsilon_{abcd} \Sigma_{\Delta}^{ab} \Sigma_{\Delta}^{cd} = 0 \quad (\text{diagonal simplicity constraint})$$

and

$$\epsilon_{abcd} \Sigma_{\Delta}^{ab} \Sigma_{\Delta'}^{cd} = 0 \quad \text{for all } \Delta, \Delta' \text{ sharing an edge (cross-simplicity constraint).}$$

These two sets of equations can be imposed at the level of each tetrahedron in the 4-simplex.

The remaining three equations (the “volume” constraints) are equivalent to the requirement that:

$$\epsilon_{abcd} \Sigma_{01}^{ab} \Sigma_{23}^{cd} = -\epsilon_{abcd} \Sigma_{02}^{ab} \Sigma_{13}^{cd} = \epsilon_{abcd} \Sigma_{03}^{ab} \Sigma_{12}^{cd} \propto V(\Sigma), \quad (4.3.10)$$

and can be imposed at the discrete level as the requirement that, for each 4-simplex:

$$\epsilon_{abcd} \Sigma_{\Delta}^{ab} \Sigma_{\Delta'}^{cd} = V \quad \text{for all } \Delta, \Delta' \text{ **not** sharing an edge (volume constraints)} \quad (4.3.11)$$

where V is defined by the above equation, and is interpreted, on the solutions of the constraints, as the volume of the 4-simplex.

An additional condition on the bivectors is usually considered, namely the “closure” constraint, which states that the sum of four bivectors corresponding to the faces of one tetrahedron is zero:

$$\sum_{\Delta \subset \mathbb{A}} \Sigma_{\Delta}^{ab} = 0. \quad (4.3.12)$$

This constraint can be understood in two ways. One can either view it as the condition that the triangles described by the variables Σ_{Δ}^{ab} close to form a tetrahedron [14], or as a consequence of the equations of motion. In a topologically trivial region such as the interior of a tetrahedron, a flat connection can be set to zero by a gauge transformation. Then using Stokes’ theorem, the integral over the equation $d\Sigma^{ab} = 0$ can be written as $\int_{\mathbb{A}} \Sigma^{ab} = 0$, which is the closure constraint. The canonical counterpart of this condition is the so-called Gauss constraint, which generates local gauge (rotation) transformations and is to be imposed on the quantum states of the theory.

The same picture appears in three spacetime dimensions, where there are no simplicity constraints and one directly deals with a $\mathfrak{su}(2)$ connection one-form ω^a and an $\mathfrak{su}(2)$ -valued (using $\mathfrak{su}(2) \simeq \mathbb{R}^3$) one-form e^a . Here the equation $de^a = 0$ is integrated over a (spacetime) triangle to give a closure constraint. The vectors associated to the edges of the triangle add up to zero, and thus have a consistent geometric interpretation as edge vectors in \mathbb{R}^3 . In this sense, an n -form with vanishing exterior derivative and appropriate internal indices can be given a geometric interpretation as describing n -simplices closing up to form an $(n+1)$ -simplex. We shall encounter another instance of this statement later on.

The closure constraint, being linear in the Σ ’s and local in each tetrahedron, is obviously easier to impose at the discrete level, and in the quantum theory, than the volume constraints. Thus it is a useful fact that it can indeed be imposed instead of them. More precisely, it can be shown [105] that the volume constraints in each 4-simplex are implied if one has enforced the diagonal and cross-diagonal simplicity constraints, *plus the closure conditions* everywhere, *i.e.* in all the tetrahedra of the

4-simplex (in general, *i.e.* for non-degenerate 4-simplices, involving tetrahedra belonging to different “time slices”). From a canonical perspective, this observation is usually phrased as an interpretation of the volume constraints as “secondary constraints” required to guarantee conservation of the other constraints (including the Gauss (closure) constraint) under time evolution.

After a period of investigations, several potentially worrying issues have been put forward regarding the Barrett-Crane model [1, 19] (for a more recent analysis of the geometry of the Barrett-Crane model, see [16]), and have given impetus to the development of alternative spin foam models [49, 57]. These models are known to have nice semiclassical properties [18], and, importantly, generalise the spin foam setting to include the Barbero-Immirzi parameter at the quantum level (for an early attempt, see [107]), and thanks to this allow for a more direct contact with the canonical loop quantum gravity. Their study is still somewhat preliminary, but the above properties make them promising candidates for a quantum theory related to gravity. One of the central features of the new models is the replacement of the quadratic simplicity constraints (4.3.9) by linear constraints of the form

$$n_a(\triangle)\Sigma^{ab}(\triangle) = 0 \quad \forall \triangle \subset \triangle, \quad (4.3.13)$$

where n_a is the normal associated to the tetrahedron \triangle and \triangle is any of the faces of \triangle .

It can be shown that these are lightly stronger than the discrete diagonal and off-diagonal quadratic simplicity constraints, and remove some of the discrete ambiguity in the solution for Σ^{ab} : out of the classes of solutions (4.3.8), one can restrict to (the discrete version of) $\Sigma^{ab} = \pm e^a \wedge e^b$ only. For a geometric analysis of these conditions in the discrete setting, see [40, 49, 57], and for a proof that the same discrete conditions can also lead to the Barrett-Crane model, see [16].

4.4 Linear Constraints for BF-Plebanski Theory

The purpose of this chapter is to investigate whether a formulation in terms of linear constraints is also possible in the classical continuum theory, and what it implies.

Let us work backwards, at first. Assume that the two-form field Σ^{ab} is of the form $\Sigma^{ab} = E^a \wedge E^b$, and that the “frame field” (not necessarily associated with a metric) E^a is non-degenerate, *i.e.* that the matrix (E_μ^a) is invertible. It follows that

$$E_a^\rho \Sigma_{\mu\nu}^{ab} = \delta_\mu^\rho E_\nu^b - \delta_\nu^\rho E_\mu^b. \quad (4.4.1)$$

In order to make a connection to the discrete setting it is more convenient to work with exterior powers of the cotangent bundle only (n -forms can be integrated over n -dimensional submanifolds). Hence we multiply (4.4.1) by $\epsilon_{\rho\sigma\tau\nu}$ and insert the relation $\epsilon_{\rho\sigma\tau\nu}E_a^\rho = (\det E_a^\mu)\epsilon_{\mu\sigma\tau\nu}E_\sigma^dE_\tau^eE_\nu^f$, which is true for invertible matrices, obtaining

$$\epsilon_{\mu\sigma\tau\nu}E_\sigma^dE_\tau^eE_\nu^f\Sigma_{\mu\nu}^{ab} = (\det E_\mu^a)(\epsilon_{\mu\sigma\tau\nu}E_\nu^b - \epsilon_{\nu\sigma\tau\nu}E_\mu^b). \quad (4.4.2)$$

One can define the three-form $n_{a\sigma\tau\nu} \equiv n_{a[\sigma\tau\nu]} \equiv \epsilon_{\mu\sigma\tau\nu}E_\sigma^dE_\tau^eE_\nu^f$, so that (4.4.2) take the form

$$n_{a\sigma\tau\nu}\Sigma_{\mu\nu}^{ab} = (\det E_\mu^a)(\epsilon_{\mu\sigma\tau\nu}E_\nu^b - \epsilon_{\nu\sigma\tau\nu}E_\mu^b). \quad (4.4.3)$$

$n_{a\sigma\tau\nu}$ can be interpreted as a 3D volume form for the submanifold parametrised by $(x^\sigma, x^\tau, x^\nu)$ embedded in 4D spacetime, whose internal index gives the normal to this submanifold. If E^a are a basis of one-forms at each spacetime point, then n_a are a basis of three-forms at each spacetime point, and so one can choose to work either with one or the other set of variables. Clearly E^a can be reconstructed from n_a :

$$\frac{1}{6}\epsilon^{\rho\sigma\tau\nu}n_{a\sigma\tau\nu} = s(\det E_\mu^a)E_a^\rho = s\sqrt[3]{\det\left(\frac{1}{6}\epsilon^{\nu\sigma\tau\nu}n_{b\sigma\tau\nu}\right)}E_a^\rho. \quad (4.4.4)$$

This means that the set of variables $n_a(x)$ define a co-tetrad frame at any point of the spacetime manifold (for the discrete analogue of the above, see [40]).

4.4.1 Linearised Diagonal and Off-Diagonal Constraints

So far we have just rewritten the equation we want to obtain for Σ^{ab} . Let us now consider the implications of imposing (4.4.3) as constraints, where we restrict to those with zero right-hand side, *i.e.* those for which $\{\mu, \nu\} \subset \{\sigma, \tau, \nu\}$. These are half of the equations (4.4.3). This will identify the continuum analogue of the linear simplicity constraints.

Claim 4.4.1 *For a basis n_a of three-forms, the general solution to*

$$n_{a\sigma\tau\nu}\Sigma_{\mu\nu}^{ab} = 0 \quad \forall \{\mu, \nu\} \subset \{\sigma, \tau, \nu\} \quad (4.4.5)$$

is

$$\Sigma_{\mu\nu}^{ab} = G_{\mu\nu}E_\mu^{[a}E_\nu^{b]}, \quad (4.4.6)$$

where E_μ^a is defined in terms of $n_{a\sigma\tau\nu}$ as in (4.4.4), and so in particular non-degenerate, and $G_{\mu\nu} = G_{\nu\mu}$ and $G_{\mu\mu} = 0$. (Obviously, as the variables Σ^{ab} and E^a , the “coefficients” $G_{\mu\nu}$ are spacetime dependent.)

Proof. First note that we can rewrite (4.4.5) as

$$\epsilon_{adef} E_\sigma^d E_\tau^e E_\nu^f \Sigma_{\mu\nu}^{ab} = 0 \quad (4.4.7)$$

with E_μ^a defined by (4.4.4). Then E_μ^a by assumption defines a basis in the cotangent space, so that

$$\Sigma_{\mu\nu}^{ab} = G_{\mu\nu}^{\chi\xi} E_\chi^a E_\xi^b \quad (4.4.8)$$

for some coefficients $G_{\mu\nu}^{\chi\xi}$ with $G_{\mu\nu}^{\chi\xi} \equiv G_{[\mu\nu]}^{[\chi\xi]}$. Substituting this into (4.4.7), we get

$$0 \stackrel{!}{=} \epsilon_{adef} E_\chi^a E_\sigma^d E_\tau^e E_\nu^f G_{\mu\nu}^{\chi\xi} E_\xi^b = \epsilon_{\chi\sigma\tau\nu} \det(E_\mu^a) E_\xi^b G_{\mu\nu}^{\chi\xi}, \quad (4.4.9)$$

and since $\det(E_\mu^a) \neq 0$ and E_ξ^b form a basis of (the internal) \mathbb{R}^4 , this implies that

$$\epsilon_{\chi\sigma\tau\nu} G_{\mu\nu}^{\chi\xi} = 0 \quad \forall \{\mu, \nu\} \subset \{\sigma, \tau, \nu\}. \quad (4.4.10)$$

It follows that $G_{\mu\nu}^{\chi\xi} = 0$ unless $\{\chi, \xi\} = \{\mu, \nu\}$ and so $G_{\mu\nu}^{\chi\xi} \equiv \delta_\mu^{[\chi} \delta_\nu^{\xi]} G_{\mu\nu}$.

□

By a linear redefinition $e_\mu^a = \lambda_\mu E_\mu^a$ one might try to set some of the $G_{\mu\nu}$ to a given value (usually ± 1 , but one might prefer $\pm \frac{1}{8\pi\gamma G}$), but it is clear that one needs two conditions on the $G_{\mu\nu}$ for this to be possible.

In the discrete context, one sets $n_a(\Delta) = (1, 0, 0, 0)$ for each Δ by a gauge transformation.^{||} One could use some of the gauge freedom here to restrict the form of n_a : This amounts to finding a convenient parametrisation for the coset space $GL(4)/SO(3, 1)$. Let us make the (usual) assumption that the normal to hypersurfaces $\{t = \text{constant}\}$ is indeed timelike. Then one can use the boost part of $SO(3, 1)$ to set $n_{a123} = (C, 0, 0, 0)$. The remaining $SO(3)$ subgroup can then be used to make the (3×3) matrix $n_{i0\sigma\tau}$, where $i \in \{1, 2, 3\}$, upper diagonal, so that one has the form

$$n_{a\sigma\tau\nu} \sim \begin{pmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \end{pmatrix}. \quad (4.4.11)$$

Clearly, when this form of $n_{a\sigma\tau\nu}$ is assumed, integrating the three-form n_a over a region where t is constant gives a vector in \mathbb{R}^4 that only has a time component. Its magnitude specifies the three-dimensional volume of such a region.

^{||}This seems to involve an implicit assumption, namely that there is a non-degenerate normal to each tetrahedron, as well.

4.4.2 Linearised Volume Constraints

As in the quadratic case, further constraints, in addition to the linear simplicity constraints (4.4.5), are needed to complete the identification $\Sigma^{ab} = \pm e^a \wedge e^b$.

First of all, one can show the following.

Claim 4.4.2 *Under the assumption that all $G_{\mu\nu}$ are non-zero, the necessary and sufficient conditions for the existence of a linear redefinition $e_\mu^a = \lambda_\mu E_\mu^a$, such that either $\Sigma^{ab} = ce^a \wedge e^b$ or $\Sigma^{ab} = -ce^a \wedge e^b$, where c is a given positive number, are*

$$G_{12}G_{03} = G_{01}G_{23} = G_{13}G_{02} (\neq 0). \quad (4.4.12)$$

Proof. Set $c = 1$; the extension to arbitrary c amounts to a further rescaling by \sqrt{c} . Then the required redefinition is possible if and only if there exist $\lambda_0, \dots, \lambda_3$, such that either $G_{\mu\nu} = \lambda_\mu \lambda_\nu$ for all $\mu \neq \nu$ or $G_{\mu\nu} = -\lambda_\mu \lambda_\nu$ for all $\mu \neq \nu$. Clearly (4.4.12) are necessary. They are also sufficient: Take

$$\lambda_1 = \sqrt{\left| \frac{G_{12}G_{13}}{G_{23}} \right|}, \quad \lambda_2 = \operatorname{sgn} \left(\frac{G_{12}G_{13}}{G_{23}} \right) \frac{G_{12}}{\lambda_1}, \quad \lambda_3 = \operatorname{sgn} \left(\frac{G_{12}G_{13}}{G_{23}} \right) \frac{G_{13}}{\lambda_1}, \quad (4.4.13)$$

which solves the equations for G_{12}, G_{13} and G_{23} with $\operatorname{sgn} \left(\frac{G_{12}G_{13}}{G_{23}} \right)$ specifying the overall sign. The remaining three equations for G_{01}, G_{02} and G_{03} are then solved by the two relations (4.4.12) and

$$\lambda_0 = \operatorname{sgn} \left(\frac{G_{12}G_{13}}{G_{23}} \right) \frac{G_{01}}{\lambda_1}. \quad (4.4.14)$$

□

The assumption $G_{\mu\nu} \neq 0$ is necessary: One solution to (4.4.12) is $G_{12} = G_{23} = G_{13} = 0$ with the other $G_{\mu\nu}$ non-zero, which cannot be expressed as $G_{\mu\nu} = \pm \lambda_\mu \lambda_\nu$.

Further constraints, in addition to the linear simplicity constraints (4.4.5), are needed to complete the identification $\Sigma^{ab} = \pm e^a \wedge e^b$. One possibility is to use the quadratic volume constraints (4.4.12). Take the three volume constraints (4.3.10),

$$\epsilon_{abcd} \Sigma_{01}^{ab} \Sigma_{23}^{cd} = -\epsilon_{abcd} \Sigma_{02}^{ab} \Sigma_{13}^{cd} = \epsilon_{abcd} \Sigma_{03}^{ab} \Sigma_{12}^{cd}, \quad (4.4.15)$$

and substitute the solution $\Sigma_{\mu\nu}^{ab} = G_{\mu\nu} E_\mu^{[a} E_\nu^{b]}$ of (4.4.5). This gives precisely (4.4.12). The non-degeneracy assumption needed for (4.4.12) is then the usual one, namely $V \neq 0$ in (4.3.6).

This shows that imposing the linear version of the diagonal and off-diagonal simplicity constraints (4.4.5) together with the quadratic volume constraints (4.3.10) and a non-degeneracy assumption on Σ^{ab} implies that

$$\Sigma^{ab} = \pm c e^a \wedge e^b \quad (4.4.16)$$

for some set of one-forms e^a , where $c > 0$ can be chosen at will. Thus, linearising the diagonal and off-diagonal simplicity constraints means that two of the four types of solutions for Σ^{ab} are removed, but on the other hand one needs to introduce a basis of three-forms n_a at each spacetime point, which is put in as an additional variable. One also still has to assume $V \neq 0$.

There is also generically no evolution of initial data with $V \neq 0$ into a degenerate Σ^{ab} with $V = 0$ and a non-geometric interpretation (this is part of the discussion of [143]). The geometry of the spacetime manifold is specified by e^a and not by E^a which is only used to determine normals in the constraints.

Alternatively, one might prefer to use a linear version of the volume constraints as well. Consider the original equation (4.4.3)

$$n_{a\sigma\tau\nu}\Sigma_{\mu\nu}^{ab} = (\det E_\mu^a) (\epsilon_{\mu\sigma\tau\nu}E_\nu^b - \epsilon_{\nu\sigma\tau\nu}E_\mu^b), \quad (4.4.17)$$

which was equivalent to $\Sigma^{ab} = E^a \wedge E^b$ for a basis of one-forms E^a . So far we only considered one half of these equations, namely those with $\{\mu, \nu\} \subset \{\sigma, \tau, \nu\}$. The other half have the form

$$n_{a\nu\tau\nu}\Sigma_{\mu\nu}^{ab} = (\det E_\mu^a)\epsilon_{\mu\nu\tau\nu}E_\nu^b, \quad \text{no sum over } \nu, \quad (4.4.18)$$

with $\epsilon_{\mu\nu\tau\nu} \neq 0$. One way to read these equations is as the requirement on the left-hand side to be totally antisymmetric in (μ, τ, ν) :

$$n_{a\nu\tau\nu}\Sigma_{\mu\nu}^{ab} = n_{a\nu\nu\mu}\Sigma_{\tau\nu}^{ab} = n_{a\nu\mu\tau}\Sigma_{\nu\nu}^{ab}, \quad \text{no sum over } \nu. \quad (4.4.19)$$

We could again try to turn the argument around and impose (4.4.19) as constraints on a \mathfrak{g} -valued two-form Σ^{ab} together with the linear simplicity constraints (4.4.5). Substituting the solution $\Sigma_{\mu\nu}^{ab} = G_{\mu\nu}E_\mu^{[a}E_\nu^{b]}$ of the linear simplicity constraints into (4.4.19) gives (after diving by a non-zero factor $\frac{1}{2}\det(E_\mu^a)$)

$$\epsilon_{\mu\nu\tau\nu}G_{\mu\nu}E_\nu^b = \epsilon_{\tau\nu\nu\mu}G_{\tau\nu}E_\nu^b = \epsilon_{\nu\nu\mu\tau}G_{\nu\nu}E_\nu^b. \quad (4.4.20)$$

For $\epsilon_{\mu\nu\tau\nu} \neq 0$ this would imply $G_{\mu\nu} = G_{\tau\nu} = G_{v\nu}$. By Claim 4.4.2, imposing (4.4.19) for one fixed ν , say $\nu = 0$, is generically not sufficient: If we know that $G_{01} = G_{02} = G_{03} \neq 0$, we still have the condition

$$G_{12} = G_{13} = G_{23}, \quad (4.4.21)$$

so that one needs more conditions of the form (4.4.19). These will then imply that all $G_{\mu\nu}$ are equal, $G_{\mu\nu} = c'$ for some c' that could be positive, negative, or zero. One can absorb $|c'|$ by an overall redefinition, so that one has

$$\Sigma_{ab} = \pm c e^a \wedge e^b, \quad (4.4.22)$$

for any chosen c , as before. Note that here it is possible, if $c' = 0$ at a point, that all e^a are zero this point and so $\Sigma^{ab} = 0$ as well. While this is a very degenerate geometry, it is still a geometry.

While the conditions (4.4.19), imposed for all values of ν , are therefore sufficient to complete the identification of the two-form field Σ^{ab} as $\pm c e^a \wedge e^b$, note that (4.4.19) is a massively redundant set of constraints: In order to obtain at most five relations on the coefficients $G_{\mu\nu}$ (two relations (4.4.12) if all $G_{\mu\nu}$ are nonzero), we are imposing *eight vector* equations! We have not exploited the fact that (4.4.20) is a multiple of one of the vectors E_ν^b , which are by assumption linearly independent. We could add several of the equations (4.4.19) for different ν , instead of considering all equations for different ν separately. Let us try to impose

$$\sum_{\nu} \sum_{\{\mu, v\} \notin \{\nu, \tau\}} n_{\mu\nu\tau\nu} \Sigma_{\mu\nu}^{ab} = 0, \quad \tau \in \{0, 1, 2, 3\} \text{ fixed.} \quad (4.4.23)$$

Again substituting the solution $\Sigma_{\mu\nu}^{ab} = G_{\mu\nu} E_\mu^{[a} E_\nu^{b]}$ of the linear simplicity constraints into (4.4.23), we obtain

$$\frac{1}{2} \det(E_\mu^a) \sum_{\nu} \sum_{\{\mu, v\} \notin \{\nu, \tau\}} \epsilon_{\mu\nu\tau\nu} G_{\mu\nu} E_\nu^b = 0, \quad \tau \text{ fixed,} \quad (4.4.24)$$

which implies, by linear independence of the E_ν^b , that indeed $G_{\mu\nu} = G_{v\nu}$ for all $\tau \notin \{\mu, \nu, v\}$. It is then sufficient to impose the constraint (4.4.23) for three different choices of τ , say $\tau = 0, 1, 2$, so that we only need three vector equations instead of eight.

By absorbing the constant $G_{\mu\nu} = c'$ (all $G_{\mu\nu}$ are equal) we rescale all E^a by the same factor to obtain the variables e^a that will have the physical interpretation of frame

fields encoding the metric geometry of spacetime. While in the case of quadratic volume constraints the one-forms E^a , or alternatively the three-forms n^a , only specified the normals to submanifolds $\{x^\mu = \text{constant}\}$, for linear volume constraints they can be directly interpreted, up to a position-dependent normalisation, as specifying an orthonormal basis in the cotangent space.

Note that this implies that one can assume a convenient normalisation for the one-forms E^a . Instead of just assuming non-degeneracy $\det(E_\mu^a) \neq 0$, one could fix $\det(E_\mu^a) = 1$. This is no restriction of the physical content of the theory as the E^a , for both linear and quadratic volume constraints, only have a geometric interpretation after rescaling. One could then interpret E_μ^a as a map into $SL(4, \mathbb{R})$. For linear volume constraints, the relation between the normalised one-forms E^a and the variables e^a that are interpreted as frame fields is a single function on spacetime which may be viewed as a “gauge” in the sense of Weyl [163].

In contrast to the case of the quadratic volume constraint, no non-degeneracy assumption on the two-form Σ^{ab} is needed to enforce simplicity. One might get $\Sigma^{ab} = 0$ in some region as a solution to the constraints, in which case the action for this region will be zero. This is analogous to a metric with vanishing determinant in general relativity and, in contrast to the requirement $V \neq 0$ outlined above, not an additional issue. Notice, however, that one still has to assume that the tetrad field E^a and, equivalently, the co-tetrad field n_a are non-degenerate, in order for the simplicity and volume constraints to imply (4.4.16). Failing this, one gets solutions of the constraints that admit no proper geometric interpretation.

In the end, writing the action for BF theory in terms of Σ^{ab} ,

$$S = \int B^{ab} \wedge R_{ab} = \int \Sigma^{ab} \wedge R_{ab} + \frac{\gamma}{2} \epsilon^{ab}_{cd} \Sigma^{cd} \wedge R_{ab}, \quad (4.4.25)$$

we substitute (4.4.16) into this action, which gives (setting $c = \frac{1}{8\pi\gamma G}$)

$$S = \frac{1}{8\pi G} \int_{\mathcal{R} \times \mathbb{R}} \sigma(x) \left(\frac{1}{2} \epsilon_{abcd} e^a \wedge e^b \wedge R^{cd} + \frac{1}{\gamma} e^a \wedge e^b \wedge R_{ab} \right). \quad (4.4.26)$$

One is left with a field $\sigma(x)$ that can take the values ± 1 , but in the classical theory one may again argue that if $\sigma = 1$ everywhere on an initial hypersurface, there will be no evolution into $\sigma = -1$. What we obtain is first order general relativity where one uses $(\det e) = \pm |\det e|$ instead of $|\det e|$ as a volume element in the action. If σ is continuous as classical fields usually are assumed to be, this differs from the action with $|\det e|$ by an overall sign at most.

To summarise, we have identified both a linear version of the quadratic simplicity constraints and a linear version of the (quadratic) volume constraints in the continuum, which can be used to reduce topological BF theory to 4D gravity in the continuum. We have found also that both linear versions are slightly stronger (*i.e.* more restrictive) than the corresponding quadratic constraints, so that the resulting constrained theory is likely to be closer to gravity at the quantum level than the one in which quadratic constraints are implemented. We now discuss the discrete counterpart of the constraints found above.

4.4.3 Discrete Linear Constraints and their Relations

The discrete analogue of (4.4.5) is just the linear constraint used in [49, 57], as desired:

$$n_a(\triangle)\Sigma^{ab}(\triangle) = 0 \quad \forall \triangle \subset \triangle. \quad (4.4.27)$$

One could write down also a discrete version of (4.4.19), obtained in the natural way, demanding that within the same 4-simplex

$$n_a(\triangle)\Sigma^{ab}(\triangle') = n_a(\triangle')\Sigma^{ab}(\triangle'') \quad (4.4.28)$$

whenever $\triangle' \not\subset \triangle$ and $\triangle'' \not\subset \triangle'$ and the edge shared by \triangle' and \triangle is the same as that shared by \triangle' and \triangle'' .

In the following we adopt the notation of [105], where the tetrahedra in a given 4-simplex are labelled by **A, B, C, D, E**, so that triangles are represented by **AB, AC**, etc., and edges by combinations **ABC, ABD**, etc. The orientation of the triangles and tetrahedra in (4.4.28) is then fixed by the signs of the permutations of the letters,

$$n_a(\triangle_{\mathbf{A}})\Sigma^{ab}(\triangle_{\mathbf{BC}}) = -n_a(\triangle_{\mathbf{B}})\Sigma^{ab}(\triangle_{\mathbf{AC}}) = n_a(\triangle_{\mathbf{C}})\Sigma^{ab}(\triangle_{\mathbf{AB}}), \quad \text{etc.} \quad (4.4.29)$$

In analogy to the continuum case, it will be sufficient to impose, instead of the full set of conditions (4.4.29), certain linear combinations of (4.4.29) to complete the geometric interpretation of the bivectors $\Sigma^{ab}(\triangle)$. The discrete analogue of the three continuum equations (4.4.23), where the index e was kept fixed, is to pick one of the tetrahedra and add those six equations out of (4.4.29) which involve triangles belonging to this tetrahedron. Starting with **A**, we impose the constraint

$$\sum_{\{i,j\} \not\supset A} n_a(\triangle_i)\Sigma^{ab}(\triangle_{\mathbf{Aj}}) = 0, \quad (4.4.30)$$

and the equivalent conditions for the tetrahedra \mathbf{B} to \mathbf{E} , thereby needing to satisfy only five instead of 20 volume constraints.

The above discrete formulation of the linearised volume constraints resembles strongly the edge simplicity constraints studied, in a canonical setting, in [47], and it imposes indeed the same restriction on the discrete data. However, it does not match exactly any of the various expressions given for these edge simplicity constraints in [47]. The correspondence between the two, therefore, deserves to be studied in more detail, given also that edge simplicity constraints have been shown to be crucial for the kinematical phase space of BF theory (and of loop gravity) to reduce to that of discrete gravity, in accordance with what we find here in a covariant setting.

In spin foam models such as [49, 57], as we mentioned earlier, only the diagonal and off-diagonal simplicity constraints, but no quadratic volume constraints (4.3.10) are imposed. This is because in the discrete setting, one can use the closure constraint (4.3.12), imposed in all the tetrahedra in a 4-simplex, to relate the (quadratic) simplicity constraints to the volume constraints, so that if the former are imposed everywhere the latter follows. Since the quadratic simplicity constraints follow from the linear ones, as can be easily checked, this argument is still valid if one uses linear simplicity constraints.

One might hope that the sufficient set of linear volume constraints (4.4.30) would also follow from the linear simplicity constraints and the closure constraints. This is almost the case, but not quite. In fact, one more constraint should be added to simplicity and closure imposed in the five tetrahedra in the 4-simplex. This is a “4D closure” constraint of the form

$$n_a(\triangle_{\mathbf{A}}) + n_a(\triangle_{\mathbf{B}}) + n_a(\triangle_{\mathbf{C}}) + n_a(\triangle_{\mathbf{D}}) + n_a(\triangle_{\mathbf{E}}) = 0, \quad (4.4.31)$$

where \triangle_i are the (appropriately oriented) tetrahedra of a given 4-simplex.

Just as for the usual closure constraint (4.3.12), there are two ways to understand why such a constraint must be imposed. Recall that if one demands the triangles described by discrete variables Σ_{\triangle}^{ab} close to form a tetrahedron, they have to satisfy (4.3.12). Alternatively, one can start with the continuum field equation $\nabla_{[\mu}^{(\omega)} \Sigma_{\nu\rho]}^{ab} = 0$, where $\nabla^{(\omega)}$ is the covariant derivative for the connection ω^{ab} , set the (flat) connection to zero by a gauge transformation, and integrate this over an infinitesimal 3-ball (whose triangulation is a tetrahedron).

The new constraint (4.4.31) seems to be the analogous statement that tetrahedra close up to form a 4-simplex. By Hodge duality $\wedge^1 \mathbb{R}^4 \simeq \wedge^3 \mathbb{R}^4$ and any internal cov-

ector $n_a(\triangle)$ can be mapped to a three-form; unlike for two-forms, any three-form can be written as $E^1 \wedge E^2 \wedge E^3$ for some e^α . Demanding that the tetrahedra described by these three-forms form a closed surface is then (4.4.31). Thus the simplicial geometric reasoning goes through also for this new constraint. In terms of the equations of motion of the theory, on the other hand, the only argument for the need of this constraint is the following. If $\nabla_{[\mu}^{(\omega)} \Sigma_{\nu\rho]}^{ab} = 0$ and we assume that $\Sigma^{ab} = \pm e^a \wedge e^b$, then it follows that $\nabla_{[\mu}^{(\omega)} n_{\nu\rho\sigma]}^a = 0$ as well. Integrating this equation (with the connection again set to zero) over a 4-ball (which can be thought of as our 4-simplex) whose boundary is a 3-sphere, triangulated by tetrahedra, then leads to (4.4.31). We then however have to assume simplicity of Σ^{ab} . A more direct derivation of (4.4.31) from the equations of motion would be desirable.

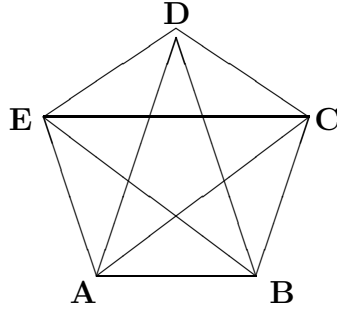


Figure 4.3: Dual picture of a 4-simplex; points represent tetrahedra, lines represent triangles shared by adjacent tetrahedra.

The role of this constraint, anyway, is the following. Consider the closure constraint

$$\Sigma^{ab}(\triangle_{\mathbf{AB}}) + \Sigma^{ab}(\triangle_{\mathbf{AC}}) + \Sigma^{ab}(\triangle_{\mathbf{AD}}) + \Sigma^{ab}(\triangle_{\mathbf{AE}}) = 0. \quad (4.4.32)$$

Contracting with $n_a(\triangle_{\mathbf{B}})$ gives, using the linear simplicity constraint $n_a(\triangle_{\mathbf{B}})\Sigma^{ab}(\triangle_{\mathbf{AB}}) = 0$,

$$n_a(\triangle_{\mathbf{B}})\Sigma^{ab}(\triangle_{\mathbf{AC}}) + n_a(\triangle_{\mathbf{B}})\Sigma^{ab}(\triangle_{\mathbf{AD}}) + n_a(\triangle_{\mathbf{B}})\Sigma^{ab}(\triangle_{\mathbf{AE}}) = 0. \quad (4.4.33)$$

Alternatively, one may start with the 4d closure constraint and contract with $\Sigma^{AB}(\triangle_{\mathbf{AB}})$ to get, again using the linear simplicity constraints,

$$n_a(\triangle_{\mathbf{C}})\Sigma^{ab}(\triangle_{\mathbf{AB}}) + n_a(\triangle_{\mathbf{D}})\Sigma^{ab}(\triangle_{\mathbf{AB}}) + n_a(\triangle_{\mathbf{E}})\Sigma^{ab}(\triangle_{\mathbf{AB}}) = 0. \quad (4.4.34)$$

In total one obtains $20 + 10 = 30$ equations of this kind that can be used to express some of the combinations $n_a(\triangle)\Sigma^{ab}(\triangle)$ in terms of others. Substituting the resulting expressions into the five discrete volume constraints (4.4.30) one finds that the

equations (4.4.30) indeed follow from the relations (4.4.33) and (4.4.34). We have seen in the continuum that the summed constraints (4.4.23) are sufficient to identify $\Sigma^{AB} = \pm e^A \wedge e^B$, and hence we find that in the discrete case the situation is analogous to the case of quadratic constraints in that a sufficient set of volume constraints can be viewed as secondary.

To see more clearly what happens in both our construction and in the case of quadratic constraints analyzed in [105], note that in our linear case one could use the 3d and 4d closure constraints to express the variables $n_a(\triangle_{\mathbf{E}})$ and $\Sigma^{ab}(\triangle_{\mathbf{AE}}), \Sigma^{ab}(\triangle_{\mathbf{BE}}), \Sigma^{ab}(\triangle_{\mathbf{CE}}), \Sigma^{ab}(\triangle_{\mathbf{DE}})$ in terms of the others. Taking the linear simplicity constraints into account, one is then left with twelve independent combinations $n_a(\triangle) \Sigma^{ab}(\triangle)$, just as in the continuum. In the continuum, we saw that one can impose the three additional constraints (4.4.23) on the twelve contractions $n_{a\mu\tau\nu} \Sigma_{\mu\nu}^{ab}$ to complete the identification $\Sigma^{ab} = \pm e^a \wedge e^b$. In the discrete case, one has the following three additional conditions coming from linear cross-simplicity constraints:

$$\begin{aligned} 0 &= n_a(\triangle_{\mathbf{E}}) \Sigma^{ab}(\triangle_{\mathbf{AE}}) \\ &= n_a(\triangle_{\mathbf{B}}) \Sigma^{ab}(\triangle_{\mathbf{AC}}) + n_a(\triangle_{\mathbf{B}}) \Sigma^{ab}(\triangle_{\mathbf{AD}}) + n_a(\triangle_{\mathbf{C}}) \Sigma^{ab}(\triangle_{\mathbf{AB}}) \\ &\quad + n_a(\triangle_{\mathbf{C}}) \Sigma^{ab}(\triangle_{\mathbf{AD}}) + n_a(\triangle_{\mathbf{D}}) \Sigma^{ab}(\triangle_{\mathbf{AB}}) + n_a(\triangle_{\mathbf{D}}) \Sigma^{ab}(\triangle_{\mathbf{AC}}) \end{aligned} \quad (4.4.35)$$

and similar ones coming from $n_a(\triangle_{\mathbf{E}}) \Sigma^{ab}(\triangle_{\mathbf{BE}}) = 0$ and $n_a(\triangle_{\mathbf{E}}) \Sigma^{ab}(\triangle_{\mathbf{CE}}) = 0$. These are precisely the analogue of the continuum constraints (4.4.23).

Similarly, in the case of quadratic simplicity constraints, one can use 3d closure to eliminate $\Sigma^{ab}(\triangle_{\mathbf{AE}}), \Sigma^{ab}(\triangle_{\mathbf{BE}}), \Sigma^{ab}(\triangle_{\mathbf{CE}}), \Sigma^{ab}(\triangle_{\mathbf{DE}})$. Then one observes that additional quadratic cross-simplicity constraints give expressions such as

$$\begin{aligned} 0 &= \epsilon_{abcd} \Sigma^{ab}(\triangle_{\mathbf{AE}}) \Sigma^{cd}(\triangle_{\mathbf{BE}}) \\ &= \epsilon_{abcd} \Sigma^{ab}(\triangle_{\mathbf{AC}}) \Sigma^{cd}(\triangle_{\mathbf{BD}}) + \epsilon_{abcd} \Sigma^{ab}(\triangle_{\mathbf{AD}}) \Sigma^{cd}(\triangle_{\mathbf{BC}}) \end{aligned} \quad (4.4.36)$$

which are equivalent to the desired (two) volume constraints.

All of this is an exercise in solving a system of linear equations for which there might be a more simple and elegant description, but the upshot is the following. The sufficient set of linear volume constraints (4.4.23) does indeed follow from the linear simplicity constraints and the closure constraints, once one also imposes a four-dimensional closure constraint on the normals to tetrahedra that seems very natural in light of their geometric interpretation. Just as in the formulation in terms of quadratic simplicity constraints [105], the volume constraints can be viewed as

secondary constraints that imply conservation of the simplicity constraints in time, or put differently, the volume constraints follow if the simplicity constraints hold everywhere. Once more this strenghtens the relationship between the discrete linear volume constraints we have identified and the edge simplicity constraints of [47].

4.5 Lagrangian and Hamiltonian Formulation

Let us briefly outline the Lagrangian formulation of 4D gravity resulting from our linear constraints added to BF theory. One adds the linear simplicity and volume constraints to the action of BF theory using Lagrange multipliers:

$$S = \int d^4x \left(\frac{1}{4} \epsilon^{\mu\nu\rho\sigma} \Sigma_{\mu\nu}^{ab} R_{ab\rho\sigma}[\omega] + \frac{\gamma}{8} \epsilon^{\mu\nu\rho\sigma} \epsilon_{abcd} \Sigma_{\mu\nu}^{ab} R_{\rho\sigma}^{cd}[\omega] + \Xi_b^{\mu\nu\sigma\tau\nu} n_{a\sigma\tau\nu} \Sigma_{\mu\nu}^{ab} \right), \quad (4.5.1)$$

where the Lagrange multiplier field $\Xi_b^{\mu\nu\sigma\tau\nu}$ satisfies $\Xi_b^{\mu\nu\sigma\tau\nu} \equiv \Xi_b^{[\mu\nu][\sigma\tau\nu]}$, and

$$\epsilon_{\mu\tau\nu} \Xi_b^{\mu\nu\tau\nu} = 0 \quad (\text{no sum over } \nu). \quad (4.5.2)$$

Indeed, varying with respect to Ξ_b then gives back the constraints

$$n_{a\sigma\tau\nu} \Sigma_{\mu\nu}^{ab} = \begin{cases} 0, & \{\mu, \nu\} \subset \{\sigma, \tau, \nu\}, \\ \epsilon_{\mu\tau\nu} f_\nu^b, & \nu = \sigma \text{ (for some } f_\nu^b \text{)}. \end{cases} \quad (4.5.3)$$

Note that the second line corresponds to the set of constraints (4.4.19) and not to the summed version (4.4.23), and that it is clearly sufficient for the geometric interpretation of Σ^{ab} . The field equation from varying with respect to the connection ω is the usual

$$\nabla_{[\mu}^{(\omega)} \Sigma_{\nu\rho]}^{ab} = 0, \quad (4.5.4)$$

where ∇ is the covariant derivative for the connection ω^{ab} . The remaining equations involve the Lagrange multipliers, as would be expected:

$$\frac{1}{4} \epsilon^{\mu\nu\rho\sigma} R_{ab\rho\sigma}[\omega] + \frac{\gamma}{8} \epsilon^{\mu\nu\rho\sigma} \epsilon_{abcd} R_{\rho\sigma}^{cd}[\omega] + \Xi_b^{\mu\nu\sigma\tau\nu} n_{a\sigma\tau\nu} = 0, \quad \Xi_b^{\mu\nu\sigma\tau\nu} \Sigma_{\mu\nu}^{ab} = 0. \quad (4.5.5)$$

We have seen that the constraints imply that $\Sigma^{ab} = \pm e^a \wedge e^b$, and when substituting this back into the action one will recover general relativity, modulo the possible sign ambiguity we have already discussed.

We leave a complete Hamiltonian analysis of this theory to future work. However, we note a feature of the theory that follows directly from the use of linear constraints, and from the introduction of the additional variables n_a .

As in unconstrained BF theory the initial dynamical variables will be the spatial part of the connection ω_k^{ab} and its conjugate momentum $P_{ab}^k \equiv \frac{1}{2}\epsilon^{ijk}\Sigma_{abij}$. We also saw that the equation of motion $\nabla_{[\mu}^{(\omega)}\Sigma_{\nu\rho]}^{ab} = 0$ is unaffected by the constraints. Hence there will be Gauss constraints of the form

$$\mathcal{G}^{cd} \equiv \partial_i P^{cdi} + \omega^c_{ei} P^{edi} + \omega^d_{ei} P^{cei} \quad (4.5.6)$$

on the canonical momenta. Their role is to generate G gauge transformations.

Looking at the action (4.5.1), one would already require that (4.5.6) should be modified to generate gauge transformations on the normals $n_{a\sigma\tau\nu}$; (4.5.1) is only invariant under gauge transformations if the three-forms $n_{a\sigma\tau\nu}$ are transformed. The need for such a modification is also seen if one computes Poisson brackets between the linear simplicity and Gauss constraints. Define “smeared” constraints

$$C[\Xi] := \int \Xi_b^{ij,\sigma\tau\nu} n_{a\sigma\tau\nu} \epsilon_{ijk} P^{abk}, \quad \mathcal{G}[\Lambda] := \int \Lambda^{cd} \mathcal{G}_{cd}. \quad (4.5.7)$$

One then finds that

$$\begin{aligned} \{C[\Xi], \mathcal{G}[\Lambda]\} &= - \int \frac{\delta C[\Xi]}{\delta P^{abk}} \frac{\delta \mathcal{G}[\Lambda]}{\delta \omega_{abk}} \\ &= - \int \Xi_B^{ij,\sigma\tau\nu} n_{a\sigma\tau\nu} \epsilon_{ijk} \left[\Lambda^{ad} P_d^b{}^k - \Lambda^{bd} P_d^a{}^k \right] \\ &= -C[\Lambda \cdot \Xi] - \int \Xi_B^{ij,\sigma\tau\nu} \Lambda^{ad} n_{a\sigma\tau\nu} \epsilon_{ijk} P_d^b{}^k, \end{aligned} \quad (4.5.8)$$

where $(\Lambda \cdot \Xi)_d^{ij,\sigma\tau\nu} = \Lambda_d^b \Xi_b^{ij,\sigma\tau\nu}$. The first term alone would imply that $\mathcal{G}[\Lambda]$ generates gauge transformations, but the second term is an unwanted extra piece. For \mathcal{G} to be a generator of gauge transformations, it must be first class (*i.e.* commute with other constraints up to linear combinations of constraints). We can remedy this by adding the variables $n_{a\mu\nu\rho}$ to the phase space, together with their conjugate momenta $\pi^{a\mu\nu\rho}$. Now we can define a new Gauss constraint

$$\mathcal{G}'^{cd} \equiv \mathcal{G}^{cd} - n_{\mu\nu\rho}^{[c} \pi^{d]\mu\nu\rho}. \quad (4.5.9)$$

Then, computing the Poisson brackets of the new Gauss constraint with $C[\Xi]$, one finds

$$\{C[\Xi], \mathcal{G}'[\Lambda]\} = \{C[\Xi], \mathcal{G}[\Lambda]\} - \int \Xi_B^{ij,\sigma\tau\nu} \Lambda^{ca} n_{c\sigma\tau\nu} \epsilon_{ijk} P^{abk} = -C[\Lambda \cdot \Xi], \quad (4.5.10)$$

as desired. We have however increased the number of phase space variables at each point by 32.

A similar reformulation of the Gauss constraint, leading to a relaxation of the gauge invariance properties of spin network states, has been already suggested by the Hamiltonian analysis of the Plebanski theory [28], and it has been advocated in the loop quantum gravity context in [3, 106] as well as spin foam and group field theory context [2, 16, 123].

4.6 Summary and Outlook

We have investigated a formulation of classical BF-Plebanski theory where the constraint $\Sigma^{ab} = \pm e^a \wedge e^b$, needed to reproduce general relativity in four dimensions, starting from topological BF theory, is imposed through constraints linear in the bivector field Σ^{ab} . The discrete counterpart of a part of these linear constraints (the “simplicity constraints”), in fact, has proven very useful in the spin foam approach to quantum gravity [49, 57].

The corresponding continuum constraints have been easily identified, and can indeed be used to replace the quadratic “diagonal” and “off-diagonal” parts of the simplicity constraints appearing in the Plebanski formulation. As in the discrete case, one needed to introduce a new set of variables n_a which are assumed to form a basis of three-forms at each point of spacetime, and are slightly stronger than the quadratic constraints: they eliminate two of the four sectors of solutions that are present for quadratic constraints.

In the second part of the analysis we found that the quadratic volume constraints of the Plebanski formulation, needed to complete the identification $\Sigma^{ab} = \pm e^a \wedge e^b$, can also be replaced by linear constraints, which again are stronger than their quadratic analogues. They do not require an additional non-degeneracy assumption on Σ^{ab} . However, a non-degeneracy assumption on the three-forms n_a is still necessary, and only when this is imposed one can hope to eliminate all “non-geometric” degenerate configurations for Σ^{ab} , which are feared to dominate the quantum theory in the case of quadratic volume constraints. Also, while for quadratic volume constraints the variables n_a merely specify normals to submanifolds $\{x^\mu = \text{constant}\}$ and hence can be independently rescaled arbitrarily at each point, for linear volume constraints they directly specify, up to an overall rescaling, the frame field encoding the metric geometry, *i.e.* an orthonormal basis in the cotangent space at each spacetime point.

We have then analysed the discrete (simplicial) translation of the linear constraints

we identified. In the context of spin foams, the quadratic volume constraints follow from imposing the (quadratic) diagonal and off-diagonal simplicity constraints everywhere together with closure constraints on the discrete variables $\Sigma^{ab}(\Delta)$. We have shown a similar property for the linear volume constraints. If (linear) diagonal and off-diagonal simplicity constraints and closure constraints for *both* bivector variables $\Sigma^{ab}(\Delta)$ and normals $n_a(\Delta)$ are imposed everywhere, a sufficient set of linear volume constraints follows. This means that “non-geometric” bivector configurations cannot appear if the additional closure constraint on the normals holds, and the same normals are assumed to be non-degenerate.

We have not performed a complete Hamiltonian analysis of the resulting linear constrained BF action for gravity, but only noted that the use of linear simplicity and volume constraints immediately requires a modification of the usual Gauss constraint to generate a transformation of the normal 3-form variables n_a alongside that of the Σ ’s; a similar relaxation of the Gauss constraint, which translates at the spin foam and discrete gravity level into a closure constraint for simplices, and in the canonical quantum gravity context into a generalisation of spin network states, has been suggested on more than one occasion in the literature [2, 3, 16, 106], even if its proper implementation at the quantum level has not been yet developed. On the classical level, therefore, a full Hamiltonian analysis of the constraints would be highly desirable. This would involve adding momenta for the components Σ_{0i}^{ab} , which are Lagrange multipliers in unconstrained BF theory, as well as those for the normals n_a we have introduced, so that all variables can transform nontrivially under G gauge transformations generated by a modified Gauss constraint, as shown.

Still at the classical level, but with obvious implications for the quantisation, one aspect of our construction that deserves further work is the relation between the discretised linear volume constraints we have found and the edge simplicity constraints used in [47]. As noted, the two sets of constraints appear to be very similar, and their role in the classical theory is the same, in particular, they remove (partly) the non-geometric configurations from the configuration space (or phase space) of the theory and appear as “secondary” in the sense specified above. So it is natural to conjecture that one is simply a reformulation of the other. The implications for the quantum theory are not only due to the dominant role that non-geometric configurations may play in the quantum theory, if not removed, but also in the fact that one discrete formulation of these constraints can actually be simpler to implement in a spin foam

context than the other.

The possible use of our findings in the spin foam and group field theory context, and more generally in any quantisation based on the formulation of gravity as a constrained BF theory, are in fact most interesting. In particular, it seems to be important to explore how a closure constraint on normals could be implemented into existing spin foam models, given that we found it to be necessary for the full imposition of the geometric constraints on the variables of topological BF. A convenient setting to do so could be the GFT formulation of [16], since there the simplicial geometry and the contact with classical actions is brought to the forefront.

Part II

Gauge Symmetries and CP Violation

Chapter 5

$SU(3)$ and its Quotients

5.1 Introduction

The Lie group that will play a central role in this second part of the thesis is the special unitary group $SU(3)$. In the standard model of particle physics, this group plays two important roles: Firstly as the (**colour**) gauge group of quantum chromodynamics (QCD), the version of Yang-Mills theory describing the strong interactions, secondly as the group of **flavour** symmetry which relates different generations of quarks, famously introduced in Gell-Mann’s “Eightfold Way” [60]. We shall in the following focus on the electroweak sector of the standard model [72, 146, 161], where the gauge group is $SU(2) \times U(1)/\mathbb{Z}_2 \simeq U(2)$ [132], and possible extensions of it, where one adds another $SU(2)$ group acting on “right-handed” fields. We are interested in CP violating processes which can change quark flavour, which are again described by an element of $SU(3)$, the **Cabibbo-Kobayashi-Maskawa** (CKM) matrix. Ambiguities in the definition of the bases with respect to which this matrix is defined mean that all matrices obtained from a given element of $SU(3)$ by left or right multiplication by an element of the maximal Abelian subgroup $U(1)^2$ should be regarded as equivalent to the original matrix. Hence the space of CKM matrices is the double quotient $U(1)^2 \backslash SU(3) / U(1)^2$. In chapter 6 we will make statements about statistics of the Jarlskog invariant J , a measure of the magnitude of CP violation, which assume a choice of measure on this quotient space. In this chapter we study the geometry of $SU(3)$ and its quotients to motivate different possible choices of measure.

We will also discuss measures on the space of Hermitian and complex 3×3 matrices whose construction relies on the knowledge of measures on quotients of $SU(3)$, and

which become important if one considers the space of quark mass matrices.

Let us first introduce coordinates on $SU(3)$ that we find convenient. The Lie algebra $\mathfrak{su}(3)$ is generated by (i times) the Gell-Mann matrices

$$\begin{aligned} \lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \lambda_2 &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \\ \lambda_5 &= \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, & \lambda_6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, & \lambda_7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \\ \lambda_3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \lambda_8 &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \end{aligned} \quad (5.1.1)$$

To make the passage from $SU(3)$ to the single or double quotients straightforward, we note that any element of $SU(3)$ can be written as

$$U = A_L V A_R, \quad (5.1.2)$$

where

$$A_L = e^{\frac{i}{2}(3p-q)\lambda_3 + \frac{i\sqrt{3}}{2}(p+q)\lambda_8}, \quad A_R = e^{it\lambda_3 + i\sqrt{3}r\lambda_8}, \quad (5.1.3)$$

and

$$V = e^{ix\lambda_7} e^{-iw\lambda_3} e^{iy\lambda_5} e^{iw\lambda_3} e^{iz\lambda_2}, \quad (5.1.4)$$

where the ranges of the Euler angles x, y, z and the complex phases w, p, q, r, t are

$$0 \leq x, y, z \leq \frac{\pi}{2}, \quad 0 \leq w, p, q, r, t < 2\pi. \quad (5.1.5)$$

Explicitly, the matrix V takes the form

$$V = \begin{pmatrix} \cos y \cos z & \cos y \sin z & e^{-iw} \sin y \\ -\cos x \sin z - e^{iw} \sin x \sin y \cos z & \cos x \cos z - e^{iw} \sin x \sin y \sin z & \sin x \cos y \\ \sin x \sin z - e^{iw} \cos x \sin y \cos z & -\sin x \cos z - e^{iw} \cos x \sin y \sin z & \cos x \cos y \end{pmatrix}, \quad (5.1.6)$$

whereas the diagonal matrices A_L and A_R are

$$A_L = \text{diag}(e^{2ip}, e^{i(q-p)}, e^{-i(p+q)}), \quad A_R = \text{diag}(e^{i(r+t)}, e^{i(r-t)}, e^{-2ir}). \quad (5.1.7)$$

It should be clear the orbits of the left action of the subgroup $U(1)^2$ are parametrised by (p, q) , while the orbits of the right action are parametrised by (r, t) . Therefore one

may use (x, y, z, w) to parametrise the double quotient $U(1)^2 \backslash SU(3) / U(1)^2$. Indeed, (5.1.6) is commonly used in the literature as a parametrisation of the CKM matrix.

We may now obtain left-invariant forms on $SU(3)$ by computing the Maurer-Cartan form

$$U^{-1}dU = i \lambda_a \sigma_a. \quad (5.1.8)$$

The resulting expressions are somewhat involved, and we use the usual complex notation to slightly shorten the expressions (taking real and imaginary parts one recovers the σ_i which are all real),

$$\begin{aligned} \sigma_1 + i\sigma_2 &= ie^{2it} \omega - e^{2it} \sin y \cos(2z) (\sin w dx + \cos w \sin(2x) dq) + \frac{1}{4} e^{2it} \sin(2z) \Omega, \\ \sigma_3 &= \frac{1}{4} \cos(2z) \Omega + dt + \sin y \sin(2z) (\sin w dx + \cos w \sin(2x) dq), \\ \sigma_4 + i\sigma_5 &= ie^{i(3r+t+w)} \cos z dy - \cos y \sin z (ie^{i(3r+t)} dx + e^{i(3r+t)} \sin(2x) dq) \\ &\quad + e^{i(3r+t+w)} \cos y \sin y \cos z (3 dp - dw + \cos(2x) dq), \\ \sigma_6 + i\sigma_7 &= ie^{i(3r-t+w)} \sin z dy + \cos y \cos z (ie^{i(3r-t)} dx + e^{i(3r-t)} \sin(2x) dq) \\ &\quad + e^{i(3r-t+w)} \cos y \sin y \sin z (3 dp - dw + \cos(2x) dq), \\ \sigma_8 &= \frac{\sqrt{3}}{4} [-dp + 4 dr + dw + 2 dq \cos(2x) \cos^2 y + (3 dp - dw) \cos(2y)], \end{aligned} \quad (5.1.9)$$

where we defined the 1-forms

$$\begin{aligned} \omega &= dz + dx \cos w \sin y - 2 dq \cos x \sin w \sin x \sin y, \\ \Omega &= 3 dp + dw + dq \cos(2x) (\cos(2y) - 3) + (3 dp - dw) \cos(2y). \end{aligned} \quad (5.1.10)$$

A general left-invariant metric on the group will have the form

$$ds^2 = g^{ab} \sigma_a \sigma_b \quad (5.1.11)$$

for some constant symmetric matrix g^{ab} . Depending on the choice of g^{ab} , the resulting metric will have a certain subgroup of $SU(3)$ as isometries acting by right multiplication on $SU(3)$. The possible cases are detailed in [41]. The most symmetric case is the bi-invariant metric, invariant under $SU(3)_L \times SU(3)_R$, which results from taking g^{ab} proportional to δ^{ab} .

Since we will be mainly interested in the measure, we note that for any choice of g^{ab} the (left-invariant) *volume form* is, up to an (irrelevant) overall constant,

$$\begin{aligned} \mu_{\text{l.inv.}}^{SU(3)} &\propto \sigma_1 \wedge \sigma_2 \wedge \sigma_3 \wedge \sigma_4 \wedge \sigma_5 \wedge \sigma_6 \wedge \sigma_7 \wedge \sigma_8 \\ &= \frac{3\sqrt{3}}{2} \sin(2x) \sin y \cos^3 y \sin(2z) dx \wedge dy \wedge dz \wedge dw \wedge dp \wedge dq \wedge dr \wedge dt, \end{aligned} \quad (5.1.12)$$

which is independent of the complex phases w, p, q, r and t . Clearly left invariance fixes the volume element up to an overall constant: Pick a group element g (say the identity), then the space of 8-forms at g is one-dimensional. The volume form at any other group element is then determined to be the push-forward of the volume form at g under left translation. The volume form (5.1.12) will be the central quantity in many of the following calculations.

One could of course have started demanding right-invariance, computing the right-invariant forms from

$$dU U^{-1} = i \lambda_a \tau_a. \quad (5.1.13)$$

If the left- and right-invariant measures are identical the group is called *unimodular*. Compact and semisimple groups are unimodular. Hence we can use left and right invariance as interchangeable conditions on a natural measure on $SU(3)$.

If G is compact and semisimple, (the negative of) the Killing form is a bi-invariant Riemannian metric which is also an Einstein metric [9]. It may therefore be regarded as the most natural choice of metric on such a group. For a (complex) matrix Lie group this metric is defined by

$$ds^2 = \text{Tr}(dg dg^\dagger). \quad (5.1.14)$$

It can also be viewed as the induced metric on the group G , interpreted as a submanifold of $\mathbb{C}^{n \times n}$ with flat metric. (The same considerations are of course true for a group of real matrices.)

We will use this metric as the most natural choice later on. It does give the same measure as any left-invariant metric, as we have noted. For completeness, we give its form explicitly for $SU(3)$:

$$\begin{aligned} ds^2 &= \frac{1}{2} \text{Tr}(dU dU^\dagger) = \delta^{ab} \sigma_a \sigma_b \\ &= 3 dp^2 + dq^2 + 3 dr^2 + dt^2 + \frac{3}{2} (3 \cos(2y) - 1) dp dr \\ &\quad + 3 \cos^2 y (\cos(2z) dp dt + \cos(2x) dq dr) \\ &\quad + \frac{1}{2} [\cos(2x) \cos(2z) (\cos(2y) - 3) + 4 \sin(2x) \sin(2z) \sin y \cos w] dq dt \\ &\quad - \sin^2 y (3 dp - 3 dr + \cos(2x) dq - \cos(2z) dt) dw \\ &\quad + 2 \sin y \sin w (\sin(2z) dt dx - \sin(2x) dq dz) \\ &\quad + dx^2 + dy^2 + dz^2 + \sin^2 y dw^2 + 2 \sin y \cos w dx dz. \end{aligned} \quad (5.1.15)$$

5.2 The Homogeneous Space $SU(3)/U(1)^2$

The coset space $SU(3)/U(1)^2$ is a reductive geometry as explained in section 1.1: The splitting of $\mathfrak{su}(3)$ into an Abelian subalgebra $\mathfrak{u}(1)^2$ and the orthogonal subspace is invariant under the adjoint action of $SU(3)$. We may hence identify an appropriate subset of generators of $SU(3)$ and use the corresponding one-forms to define metrics on the homogeneous space $SU(3)/U(1)^2$.

The one-forms σ_3 and σ_8 correspond to the Gell-Mann matrices generating the Abelian subgroup $U(1)^2$. We can construct left-invariant metrics on $SU(3)/U(1)^2$ from the remaining six left-invariant forms $\sigma_1, \sigma_2, \sigma_4, \dots, \sigma_7$. The left group action of $SU(3)$ on the homogeneous space again guarantees the existence of a unique (up to a constant) left-invariant volume form, and we find that this unique volume form is

$$\begin{aligned} \mu_{\text{l.inv.}}^{SU(3)/U(1)^2} &\propto \sigma_1 \wedge \sigma_2 \wedge \sigma_4 \wedge \sigma_5 \wedge \sigma_6 \wedge \sigma_7 \\ &= \frac{3\sqrt{3}}{2} \sin(2x) \sin y \cos^3 y \sin(2z) dx \wedge dy \wedge dz \wedge dw \wedge dp \wedge dq, \end{aligned} \quad (5.2.1)$$

where we note that, due to the $SU(3)$ decomposition (5.1.2), we may parametrise the quotient $SU(3)/U(1)^2$ by the $SU(3)$ coordinates (x, y, z, w, p, q) . This is of course the same volume form as the one obtained by starting with the left-invariant volume form on $SU(3)$ and integrating over the coordinates r and t , or by just considering a slice $\{r = \text{constant}, t = \text{constant}\}$ in $SU(3)$ with the measure (5.1.12). (We shall see in the next section that this need not be true if one reduces to the double quotient $U(1)^2 \backslash SU(3)/U(1)^2$.)

There is an additional possible source of non-uniqueness of the measure, namely a possible ambiguity of the definition of the same manifold as different homogeneous spaces $X = G/H = G'/H'$. In practice, this does not seem to lead to different results for the measure.

It should be noted that the homogeneous space considered here, where the subgroup is the maximal Abelian subgroup, the “maximal torus”, of a compact group, is what is known as a *flag manifold*. Such a manifold is naturally a Kähler manifold. For any G -invariant complex structure on a flag manifold, there is a unique G -invariant Kähler-Einstein metric [9]. This metric may then be regarded as the most natural choice on such a homogeneous space. An explicit construction for such metrics on spaces $SU(M)/U(1)^{M-1}$, where $U(1)^{M-1}$ is the maximal torus of $SU(M)$, is given in [136]. The case $M = 6$ might be of interest for neutrino mixing; here the analogue of the CKM matrix is the **Maki-Nakagawa-Sakata matrix** [114], usually taken to be

an element of the single quotient $U(1)^2 \backslash SU(3)$. In the see-saw mechanism one adds very heavy right-handed neutrinos, and the most general mixing matrix would be an element of $U(1)^5 \backslash SU(6)$. The construction of [136] would be the starting point of an analysis of this quotient, which is beyond the scope of this thesis.

5.3 Metrics on $U(1)^2 \backslash SU(3)/U(1)^2$

The double quotient $U(1)^2 \backslash SU(3)/U(1)^2$ does not have a natural $SU(3)$ group action, because the left and right actions of $SU(3)$ on itself do not commute with the Abelian subgroup $U(1)^2$ and hence do not induce a group action on the double quotient: If g and g' are in the same equivalence class, $[g] = [g']$, then $g = hg'h'$ for some $h, h' \in U(1)^2$. For a group action to be defined on the double quotient, one would need $[j \cdot g] = [j \cdot g']$, $j \cdot g = k(j \cdot g')k' = k(j \cdot h^{-1}gh'^{-1})k'$ for some $k, k' \in U(1)^2$. This can only be done for those j that commute with h^{-1} .

It is therefore less straightforward than in the previous cases to find a geometrically motivated measure. The simplest way of reducing the left-invariant measure on the group $SU(3)$ or the homogeneous space $SU(3)/U(1)^2$ to the double quotient would be to just integrate (5.2.1) over the coordinates p and q , which is trivial since the volume form (5.2.1) is independent of p and q . One would therefore essentially use the same measure for the double quotient.

Alternatively, one might start from a metric on either the group or the homogeneous space, reduce it in some way to the double quotient, and then determine the associated volume form, but this introduces two possible sources of ambiguity; the choice of metric and the choice of method of reduction. We will see that indeed the measure on the double quotient is vastly non-unique.

For example, one may be tempted to identify the double quotient with the submanifold $p = q = r = t = 0$ of $SU(3)$, take a symmetric metric on $SU(3)$, say the bi-invariant metric (5.1.15), and set $p = q = r = t = 0$. This approach was indeed followed by Ozsváth and Schücking [126], and leads to an appealingly simple form of the metric

$$ds^2 = dx^2 + dy^2 + dz^2 + 2 \sin y \cos w \, dx \, dz + \sin^2 y \, dw^2, \quad (5.3.1)$$

but seems not to be justified geometrically: It corresponds to a gauge choice, a choice of section in the orbits of $U(1)^2 \times U(1)^2$ over the double quotient, and introduces

symmetries (three commuting Killing vectors) into the metric which should not be present. It seems much better motivated geometrically to do a Kaluza-Klein reduction of a metric, *i.e.* to project it orthogonally to the orbits of $U(1)^2$. One writes a given metric on $SU(3)$ (say) as

$$ds^2 = h_{ij}(x) (dy^i + A^i_\mu(x) dx^\mu) (dy^j + A^j_\nu(x) dx^\nu) + \tilde{g}_{\mu\nu}(x) dx^\mu dx^\nu, \quad (5.3.2)$$

and identifies $\tilde{g}_{\mu\nu}$ as a metric on the quotient space parametrised by (x^μ) .

To understand what is going wrong in (5.3.1), we compare with the simpler case of $SU(2)$. The bi-invariant metric on $SU(2)$ is

$$ds^2 = (d\psi + \cos \theta d\phi)^2 + d\theta^2 + \sin^2 \theta d\phi^2, \quad (5.3.3)$$

where $\frac{\partial}{\partial \phi}$ generates the left action of $U(1)$ and $\frac{\partial}{\partial \psi}$ generates the right action of $U(1)$. Projecting the metric orthogonally to the orbits of right translations, *à la* Kaluza-Klein, gives the round metric

$$ds^2 = d\theta^2 + \sin^2 \theta d\phi^2 \quad (5.3.4)$$

on $SU(2)/U(1) = S^2$. By contrast, simply setting $d\psi = 0$ (the analogue of the construction of Ozsváth and Schücking) instead gives the flat metric

$$ds^2 = d\theta^2 + d\phi^2. \quad (5.3.5)$$

The round metric (5.3.4) is invariant under $SO(3)$. The flat metric (5.3.5) appears to be invariant under the Euclidean group, with $\frac{\partial}{\partial \theta}$ and $\frac{\partial}{\partial \phi}$ having the appearance of translations, but these are only local symmetries since ϕ is a periodic coordinate and θ lies in an interval.

The example of $SU(2)$ also illustrates the difference between taking the left-invariant measure on the homogeneous space and a measure obtained by Kaluza-Klein reduction of a metric to a biquotient. The biquotient $U(1) \backslash SU(2) / U(1) = U(1) \backslash S^2$ is just an interval. Its metric becomes, after performing another Kaluza-Klein reduction of (5.3.4),

$$ds^2 = d\theta^2. \quad (5.3.6)$$

The measure would be $d\theta$, and not $d\theta \sin \theta$ as obtained by integrating a function $f(\theta)$ over the coordinate ϕ . It is apparent from this simple example that there are inequivalent ways of calculating integrals of a function on a right quotient that is invariant under the left group action; namely, one can *either* reduce the metric to

obtain a measure on the double quotient *or* take the measure on the single quotient and integrate out the left phases.

Performing a Kaluza-Klein reduction of the bi-invariant metric (5.1.15) we find a rather complicated metric on the double quotient, whose determinant gives the measure

$$\begin{aligned} \sqrt{\det \tilde{g}} \propto & \sin 2x \sin 2z \sin y \cos^2 y \cdot \left((\sin^2 2x + \sin^2 2z) \sin^2 y \right. \\ & + \frac{1}{8} (5 \cos 2y - 3) \sin^2 2x \sin^2 2z + \frac{1}{2} \sin 4x \sin 4z \sin^3 y \cos w \\ & \left. + \frac{1}{8} (3 \cos 2y - 5) \sin^2 2x \sin^2 2z \sin^2 y \cos^2 w \right)^{-\frac{1}{2}}. \end{aligned} \quad (5.3.7)$$

One might think that, starting from a more general (“squashed”) left-invariant metric on the homogeneous space $SU(3)/U(1)^2$, there might be some which after Kaluza-Klein reduction give rise to a simpler expression for the measure on the biquotient, but we have only found more complicated ones.

5.4 Hermitian and Complex Matrices

In chapter 6 we will require measures on the space of either all Hermitian or all complex 3×3 matrices which will appear as the space of quark mass matrices in the standard model of particle physics or certain extensions of it. These will involve the left- (or right-)invariant measure on the homogeneous space $SU(3)/U(1)^2$, as we will see shortly. Let us start with the space of 3×3 Hermitian matrices. It has a natural group action by $U(3)$ by conjugation, corresponding to a change of basis. We would like to determine a metric invariant under this action.

Any Hermitian matrix can be diagonalised, *i.e.* written in terms of a real diagonal and a unitary matrix:

$$M_H = U^\dagger D U, \quad (5.4.1)$$

It should be clear that U is only defined up to left multiplication by elements of $U(1)^2$; U should be regarded as an element of the homogeneous space $U(1)^2 \backslash SU(3) \simeq SU(3)/U(1)^2$.

The most symmetric measure on the space of Hermitian matrices is induced by the bi-invariant metric

$$ds^2 = \text{Tr}(dM_H dM_H) = \text{Tr}(dD dD) + 2\text{Tr}\left((dU U^\dagger D)^2 - (dU U^\dagger)^2 D^2\right) \quad (5.4.2)$$

which is clearly invariant under the action of $U(3)$ by conjugation. Define right-invariant one-forms τ_a on $SU(3)$ by

$$dU U^\dagger = i \lambda_a \tau_a, \quad (5.4.3)$$

this becomes* [with $D \equiv \text{diag}(D_1, D_2, D_3)$]

$$\begin{aligned} ds^2 &= \text{Tr}(dD \cdot dD) - 2\tau_a \tau_b \text{Tr}(\lambda_a [D, \lambda_b] D) \\ &= dD_1^2 + dD_2^2 + dD_3^2 + 2 \{ (D_1 - D_2)^2 (\tau_1^2 + \tau_2^2) + (D_1 - D_3)^2 (\tau_4^2 + \tau_5^2) \\ &\quad + (D_2 - D_3)^2 (\tau_6^2 + \tau_7^2) \}. \end{aligned} \quad (5.4.4)$$

The corresponding volume form is

$$(D_1 - D_2)^2 (D_1 - D_3)^2 (D_2 - D_3)^2 dD_1 \wedge dD_2 \wedge dD_3 \wedge \tau_1 \wedge \tau_2 \wedge \tau_4 \wedge \tau_5 \wedge \tau_6 \wedge \tau_7. \quad (5.4.5)$$

We have explained above that any left- or right-invariant measure on the coset $U(1)^2 \backslash SU(3)$ must be given by (5.2.1) and hence there is no need to compute the τ_a explicitly. We obtain a Riemannian measure

$$\begin{aligned} DM_H &:= (D_1 - D_2)^2 (D_1 - D_3)^2 (D_2 - D_3)^2 \sin(2x) \cos^3 y \sin y \sin(2z) dD_1 dD_2 dD_3 \\ &\quad \times dx dy dz dw dr dt \end{aligned} \quad (5.4.6)$$

on the space of Hermitian 3×3 matrices. Note that we use coordinates (x, y, z, w, r, t) on $U(1)^2 \backslash SU(3)$ according to the decomposition (5.1.2). The measure (5.4.6) favours strongly non-coinciding eigenvalues.

From (5.4.1), it is apparent that each Hermitian matrix with three distinct eigenvalues is associated with six different elements of $\mathbb{R}^3 \times U(1)^2 \backslash SU(3)$, related by the action of the discrete group \mathfrak{S}_3 :

$$M_H = U^\dagger D U = (U^\dagger P^{-1}) P D P^{-1} (P U) =: \tilde{U}^\dagger \tilde{D} \tilde{U}, \quad P \in \mathfrak{S}_3, \quad (5.4.7)$$

where \mathfrak{S}_3 is the symmetric group of degree 3 (the dihedral group of order 6, sometimes denoted by D_3 or D_6) which permutes the canonical basis vectors of \mathbb{R}^3 . The set of matrices with coinciding eigenvalues has zero measure and hence can be ignored in the present discussion.

Thus we need to consider the space $\mathbb{R}^3 \times (U(1)^2 \times \mathfrak{S}_3) \backslash SU(3)$ instead[†], restricting the coordinates on the quotient $U(1)^2 \backslash SU(3)$ to an appropriate range to pick one of

*Compare with the corresponding result for real matrices given in [71]

[†]Note that the space of Hermitian matrices is not topologically identified with this space; we have discarded a subset of Hermitian matrices which has zero measure for our construction.

the six matrices related by the \mathfrak{S}_3 action. We can use the fact that the \mathfrak{S}_3 action permutes the rows of an $SU(3)$ matrix to demand that the elements of the third column (cf. (5.1.6); note that absolute values will not depend on rephasing, *i.e.* right multiplication by an element of $U(1)^2$) satisfy the relation

$$|\sin y| \leq |\sin x \cos y| \leq |\cos x \cos y|, \quad (5.4.8)$$

which restricts the coordinates x and y to

$$0 \leq y \leq \arctan(\sin x), \quad 0 \leq x \leq \frac{\pi}{4}. \quad (5.4.9)$$

Using the right-invariant measure (5.2.1) on $U(1)^2 \backslash SU(3)$, we see that this region has precisely one-sixth of the total volume:

$$\frac{\int_0^{\pi/2} dz \int_0^{\pi/4} dx \int_0^{\arctan(\sin x)} dy \sin 2x \cos^3 y \sin y \sin 2z}{\int_0^{\pi/2} dz \int_0^{\pi/2} dx \int_0^{\pi/2} dy \sin 2x \cos^3 y \sin y \sin 2z} = \frac{1}{6}. \quad (5.4.10)$$

Notice that while the total volume of $(U(1)^2 \times \mathfrak{S}_3) \backslash SU(3)$ will be finite, the integrals over the \mathbb{R}^3 part will diverge; the eigenvalues of arbitrary Hermitian matrices can of course be arbitrarily large. One would have to introduce some kind of cutoff to make integrals over the space of Hermitian matrices convergent, as we will do later on.

The space of 3×3 complex matrices can be discussed along similar lines. Here, one has a left and right action by independent elements of $U(3)$. We should therefore determine a metric invariant under both of these group actions. An arbitrary complex matrix can be represented as

$$M_c = U_L^\dagger D U_R, \quad (5.4.11)$$

where U_L, U_R are unitary and $D = \text{diag}(D_1, D_2, D_3)$ is real diagonal. U_L and U_R are only defined up to simultaneous left multiplication by a diagonal unitary matrix

$$U_L \rightarrow A U_L, \quad U_R \rightarrow A U_R, \quad A \in U(1)^3 \quad (5.4.12)$$

which so that one has the correct number of real parameters, only 18 instead of 21. Since here the signs of the elements D_i can be changed by a different choice of U_L and U_R , and one can permute the D_i as for Hermitian matrices, there are additional discrete ambiguities given by elements of the group $\mathfrak{S}_3 \times \mathbb{Z}_2^3$. We will see shortly

that the measure, similar to the measure for Hermitian matrices, vanishes whenever elements of D coincide up to sign, so we can restrict to matrices with $D_1^2 \neq D_2^2 \neq D_3^2 \neq D_1^2$. Hence we can identify this subspace of the space of 3×3 complex matrices with $\mathbb{R}^3 \times (U(1)^3 \times \mathfrak{S}_3 \times \mathbb{Z}_2^3) \setminus (U(3) \times U(3)) \simeq \mathbb{R}_+^3 \times (U(1)^3 \times \mathfrak{S}_3) \setminus (U(3) \times U(3))$.

The metric we will use to determine a measure is

$$ds^2 = \text{Tr}(dM_c dM_c^\dagger), \quad (5.4.13)$$

which clearly is invariant under $M \rightarrow OMO'$ for $O, O' \in U(3)$. To evaluate this, use

$$\begin{aligned} dM_c &= -U_L^\dagger dU_L U_L^\dagger D U_R + U_L^\dagger dD U_R + U_L^\dagger D dU_R, \\ dM_c^\dagger &= -U_R^\dagger dU_R U_R^\dagger D U_L + U_R^\dagger dD U_L + U_R^\dagger D dU_L \end{aligned} \quad (5.4.14)$$

and $[D, dD] = 0$ to obtain

$$\begin{aligned} \text{Tr}(dM_c dM_c^\dagger) &= \text{Tr}(dD dD) - \text{Tr}(D^2 (dU_L U_L^\dagger)^2) - \text{Tr}(D^2 (dU_R U_R^\dagger)^2) \\ &\quad + 2 \text{Tr}(D dU_L U_L^\dagger D dU_R U_R^\dagger). \end{aligned} \quad (5.4.15)$$

We introduce right-invariant one-forms

$$dU_L U_L^\dagger = i \lambda_a \tau_L^a, \quad dU_R U_R^\dagger = i \lambda_b \tau_R^b, \quad (5.4.16)$$

where $\lambda_1, \dots, \lambda_8$ are the Gell-Mann matrices, and

$$\lambda_9 = \sqrt{\frac{2}{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (5.4.17)$$

so that $i\lambda_a$ are a basis for the Lie algebra $\mathfrak{u}(3)$. Then

$$\text{Tr}(dM_c dM_c^\dagger) = \text{Tr}(dD dD) + \text{Tr}(D^2 \lambda_{(a} \lambda_{b)}) (\tau_L^a \tau_L^b + \tau_R^a \tau_R^b) - \text{Tr}(D \lambda_a D \lambda_b) (\tau_L^a \tau_R^b + \tau_R^a \tau_L^b). \quad (5.4.18)$$

The only nonvanishing traces in (5.4.18) are

$$\begin{aligned}
\text{Tr}(D^2\lambda_1\lambda_1) &= \text{Tr}(D^2\lambda_2\lambda_2) = \text{Tr}(D^2\lambda_3\lambda_3) = \text{Tr}(D\lambda_3D\lambda_3) = D_1^2 + D_2^2, \\
\text{Tr}(D^2\lambda_{(3}\lambda_8)) &= \text{Tr}(D\lambda_{(3}D\lambda_8)) = \frac{1}{\sqrt{3}}(D_1^2 - D_2^2), \\
\text{Tr}(D^2\lambda_{(3}\lambda_9)) &= \text{Tr}(D\lambda_{(3}D\lambda_9)) = \sqrt{\frac{2}{3}}(D_1^2 - D_2^2), \\
\text{Tr}(D^2\lambda_{(8}\lambda_9)) &= \text{Tr}(D\lambda_{(8}D\lambda_9)) = \frac{\sqrt{2}}{3}(D_1^2 + D_2^2 - 2D_3^2), \\
\text{Tr}(D^2\lambda_4\lambda_4) &= \text{Tr}(D^2\lambda_5\lambda_5) = D_1^2 + D_3^2, \\
\text{Tr}(D^2\lambda_6\lambda_6) &= \text{Tr}(D^2\lambda_7\lambda_7) = D_2^2 + D_3^2, \\
\text{Tr}(D^2\lambda_8\lambda_8) &= \text{Tr}(D\lambda_8D\lambda_8) = \frac{1}{3}(D_1^2 + D_2^2 + 4D_3^2), \\
\text{Tr}(D^2\lambda_9\lambda_9) &= \text{Tr}(D\lambda_9D\lambda_9) = \frac{2}{3}(D_1^2 + D_2^2 + D_3^2), \\
\text{Tr}(D\lambda_1D\lambda_1) &= \text{Tr}(D\lambda_2D\lambda_2) = 2D_1D_2, \\
\text{Tr}(D^2\lambda_4\lambda_4) &= \text{Tr}(D^2\lambda_5\lambda_5) = 2D_1D_3, \\
\text{Tr}(D^2\lambda_6\lambda_6) &= \text{Tr}(D^2\lambda_7\lambda_7) = 2D_2D_3.
\end{aligned} \tag{5.4.19}$$

The metric can be written in the form

$$\begin{aligned}
ds^2 &= dD_1^2 + dD_2^2 + dD_3^2 + \frac{1}{2}(D_1 - D_2)^2(\tau_L^1 + \tau_R^1)^2 + \frac{1}{2}(D_1 + D_2)^2(\tau_L^1 - \tau_R^1)^2 \\
&+ \frac{1}{2}[(D_1 - D_2)^2(\tau_L^2 + \tau_R^2)^2 + (D_1 + D_2)^2(\tau_L^2 - \tau_R^2)^2 \\
&+ (D_1 - D_3)^2(\tau_L^4 + \tau_R^4)^2 + (D_1 + D_3)^2(\tau_L^4 - \tau_R^4)^2] \\
&+ \frac{1}{2}[(D_1 - D_3)^2(\tau_L^5 + \tau_R^5)^2 + (D_1 + D_3)^2(\tau_L^5 - \tau_R^5)^2 \\
&+ (D_2 - D_3)^2(\tau_L^6 + \tau_R^6)^2 + (D_2 + D_3)^2(\tau_L^6 - \tau_R^6)^2] \\
&+ \frac{1}{2}(D_2 - D_3)^2(\tau_L^7 + \tau_R^7)^2 + \frac{1}{2}(D_2 + D_3)^2(\tau_L^7 - \tau_R^7)^2 \\
&+ 2D_1^2 \left(\frac{1}{\sqrt{2}}(\tau_L^3 - \tau_R^3) + \frac{1}{\sqrt{6}}(\tau_L^8 - \tau_R^8) + \frac{1}{\sqrt{3}}(\tau_L^9 - \tau_R^9) \right)^2 \\
&+ 2D_2^2 \left(-\frac{1}{\sqrt{2}}(\tau_L^3 - \tau_R^3) + \frac{1}{\sqrt{6}}(\tau_L^8 - \tau_R^8) + \frac{1}{\sqrt{3}}(\tau_L^9 - \tau_R^9) \right)^2 \\
&+ 2D_3^2 \left(\sqrt{\frac{2}{3}}(\tau_L^8 - \tau_R^8) + \frac{1}{\sqrt{3}}(\tau_L^9 - \tau_R^9) \right)^2.
\end{aligned} \tag{5.4.20}$$

The fact that the metric only depends on $\tau_L^3 - \tau_R^3$ etc., and not on $\tau_L^3 + \tau_R^3$ etc., again reflects the $U(1)^3$ that has to be factored out. One finds that the volume form is

proportional to

$$(D_1^2 - D_2^2)^2 (D_1^2 - D_3^2)^2 (D_2^2 - D_3^2)^2 |D_1 D_2 D_3| dD_1 \wedge dD_2 \wedge dD_3 \wedge \tau_L^1 \wedge \tau_R^1 \wedge \dots \wedge (\tau_L^8 - \tau_R^8) \wedge (\tau_L^9 - \tau_R^9). \quad (5.4.21)$$

Note that this expression only depends on the absolute values of D_i , as expected. Since the range of the D_i is infinite, integration over these coordinates will give an infinity, as was the case for Hermitian matrices. Note the main difference to the measure (5.4.6) in the different powers of D_i .

Now assume that all quantities we are interested in are independent of the parameters on U_R , which will be the case in the discussion of chapter 6. Then we can integrate over these coordinates, obtaining a constant which is irrelevant in the averaging process.

We are then left with integrating over the space of possible matrices U_L , the coset $U(1)^3 \backslash U(3) = U(1)^2 \backslash SU(3)$, and the volume form is proportional to

$$(D_1^2 - D_2^2)^2 (D_1^2 - D_3^2)^2 (D_2^2 - D_3^2)^2 |D_1 D_2 D_3| dD_1 \wedge dD_2 \wedge dD_3 \wedge \tau_L^1 \wedge \tau_L^2 \wedge \tau_L^4 \wedge \tau_L^5 \wedge \tau_L^6 \wedge \tau_L^7. \quad (5.4.22)$$

In terms of the coordinates on $U(1)^2 \backslash SU(3)$ introduced in (5.1.2), the wedge product of right-invariant forms gives the usual measure on $SU(3)$, so that we finally get

$$DM_c = \prod_{i < j} (D_i^2 - D_j^2)^2 |D_1 D_2 D_3| \sin 2x \cos^3 y \sin y \sin 2z dD_1 dD_2 dD_3 dx dy dz dw dr dt. \quad (5.4.23)$$

The discrete \mathfrak{S}_3 symmetry can be taken into account just as in the previous case: We integrate only over one sixth of the homogeneous space $U(1)^2 \backslash SU(3)$, corresponding to

$$0 \leq y \leq \arctan(\sin x), \quad 0 \leq x \leq \frac{\pi}{4}. \quad (5.4.24)$$

This restriction amounts to removing unitary matrices that permute the elements of D and hence to fixing an ordering.

Chapter 6

Naturalness of CP Violation

“Well, I don’t like to get involved in these philosophical issues very much.”

Murray Gell-Mann (attributed)

6.1 Introduction and Motivation

In this chapter we will apply the mathematical theory of measures on quotients of $SU(3)$, outlined in the previous chapter, to the space of CKM matrices in the electroweak sector of the standard model in order to make statements about naturalness of the observed magnitude of CP violation in electroweak processes. We will see that the observed value is rather small when compared to the maximal possible magnitude and aim to reach a conclusion if this should be viewed as a fine-tuning problem. The task of making mathematically precise statements is simplified by the fact that $SU(3)$ is compact, and so normalisation of measures on $SU(3)$ or a quotient of it will normally not be problematic.

In modern theoretical physics one often tries to make statements about “naturalness” or “fine-tuning” of the observed values of fundamental parameters, where fine-tuning of a parameter is interpreted as an indication for incompleteness of the theory. Popular examples of fine-tuning problems include the quark mass hierarchy and the cosmological constant problem in particle physics. Since statements about naturalness are fundamentally of a statistical character, to make them mathematically precise one has to assume a well-motivated probability distribution on the parameter space relevant for the theory. A fine-tuning problem then indicates that the

probability distribution one has used should be modified by introducing new physical considerations. As an example, as long as observation was consistent with a vanishing cosmological constant, it seemed reasonable to assume that a postulated symmetry would constrain it to vanish. With more recent observations indicating that it must be taken to be very small and positive, there seems to be an issue of fine-tuning*. We note that much of the motivation to extend the standard model of particle physics is driven by such considerations, that it is by no means necessary to contemplate a “Multiverse” where all possible values of a given parameter are actually realised. We also need not consider anthropic arguments, which for the problem at hand would not give a satisfactory explanation of very weak CP violation; strong CP violation might be anthropically preferred since baryogenesis necessarily requires CP violation, and the CP violation in the electroweak sector does not seem sufficient [87].

Of course we can only ever observe and make measurements in a single Universe, and if a parameter takes a value that appears unlikely maybe this just means that an unlikely possibility is realised in our Universe. One is merely doing statistics. But one should recall **Bayes’ theorem** [20]

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(B|A)P(A)}{\sum_{A_i} P(B|A_i)P(A_i)}, \quad (6.1.1)$$

where one can take A as a hypothesis and B as an observation. In order to calculate the probability that a hypothesis follows from a given observation, one either needs *a priori* probabilities for both A and B and the conditional probability of B assuming A (which should be calculable), or *a priori* probabilities for a complete set of hypotheses $\{A_i\}$, together with all conditional probabilities of B given these hypotheses. These *a priori* probabilities are referred to as **priors**. In the second case the sum also has to be made explicit, and if it is an integral one needs to specify a measure.

In cosmology, the role of priors is particularly important as there is no possibility to confirm observations in different experimental situations, and $P(B)$ does not seem accessible. Here A would be a statement about our Universe. While it may be appropriate to consider $P(A)$ as independent of A (Laplace’s Principle of Indifference), a measure on the set of all universes has to be determined [61].

The situation considered in this chapter is similar, as “experiments” would involve different universes with different values of the fundamental physical parameters. For instance, one could try to use anthropic arguments, taking B to be the observation

*For an alternative interesting but presumably non-mainstream viewpoint, see [23]

that human beings have evolved in our Universe, to claim that the values of (in this case) CP violating CKM matrix parameters must take values close to the observed values (hypothesis A). The conditional probability $P(B|A)$ could in principle be calculated using baryogenesis etc., but then one still needs a measure on the space of possible values that A can take. Again, one is forced to pick a measure on the parameter space. We shall therefore investigate predictions for the magnitude of CP violation, measured by the Jarlskog invariant J , comparing different choices for the measure on the parameter space that we regard as natural. The observation by Kobayashi and Maskawa [98] that CP violation is only possible for at least three quark families has led to a Nobel prize, but the issue of possible fine-tuning in the magnitude of CP violation is much less understood.

It is true that, if one is talking about parameters in quantum field theory, they should really not be regarded as constants, but have an evolution with energy scale given by renormalisation group equations. However, since fine-tuning means a discrepancy of several orders of magnitude, it may well be that a fine-tuning problem is present at all energy scales. This is true for the quark mass hierarchy [170], and also for the case of CP violation: Recent numerical studies [36] indicate that J^2 does not run strongly with energy scale, but that the value at extremely high energies ($\sim 10^{15}\text{GeV}$) is merely about twice the value at low energies. It is then meaningful to talk about “naturalness” of the value of such a parameter.

Since any statements one tries to make depend very directly on the choice of measure, it is helpful if geometric considerations allow for a natural choice of probability distribution. We saw in the previous chapter that, if the parameter space is a homogeneous space G/H for G a compact Lie group and H a closed subgroup, the natural requirement on the measure determining the probability distribution is invariance under the left action of G , which leads to a unique measure (up to normalisation). Probabilities for a given function on G/H to take certain values are then well-defined.

We also saw that if, as in the case of the space of CKM matrices, one has a double quotient $H\backslash G/H$, the problem of determining a natural probability distribution is more involved than for a homogeneous space G/H . We have discussed several possible choices that we compare here. We will find that while there is no clearly preferred choice of measure, there always seems to be fine-tuning in the observed value for J , unless additional input is used.

In a second part we shall take a different approach, taking the observed values for

the quark masses into account by considering not the space of CKM matrices, but the space of mass matrices, as the fundamental parameter space. This is motivated by the observation that the mass matrices are directly linked to the Yukawa couplings and the Higgs vacuum expectation value, whereas the CKM matrix is only a derived quantity. The observed values for the quark masses will be taken as given, and a probability distribution be constructed that can reproduce these values. The choice made for this distribution will be as simple as possible in the following sense: The natural group action on the space of Hermitian mass matrices is the action of the unitary group $U(3)$ by conjugation, $M \rightarrow U M U^\dagger$. There is essentially a unique measure invariant under this action. This measure is then modified by introducing the simplest possible function that would allow for a modification of the expectation values for quark masses fitting observation. No further assumptions will be needed. We will then find that this simple choice for the measure gives an expectation value for J that is remarkably close to the observed value. Hence we will conclude that once one assumes the quark masses as given, one does not face an additional fine-tuning problem with J . This statement, while not new, is hence made precise using a geometrically motivated measure on the parameter space.

In the standard model one can restrict, without loss of generality, to Hermitian mass matrices, which is why we construct a measure on the space of 3×3 Hermitian matrices. However, in left-right symmetric extensions of the standard model, such as Pati-Salam, such an assumption can no longer be made and the mass matrices have to be regarded as arbitrary complex matrices[†]. This presumably has an effect on the statements one makes about naturalness of J ; a similarly well-motivated measure on the space of all 3×3 complex matrices might lead to very different results. We will therefore in a further calculation redo the same analysis for general complex matrices. We will again use the most symmetric measure on the space of mass matrices, here the space of general complex matrices, and modify it in the simplest possible way to incorporate the observed quark mass hierarchy. While this is a choice that could of course be made very differently, it is a simple choice that uses as few assumptions as possible, and that has worked very well for the case of Hermitian mass matrices. What we will find is that the resulting probability distribution is different, and the result is different too: The observed value for J now appears to be unnaturally large, since CP violation should be more heavily suppressed by the quark mass hierarchy.

[†]We thank Ben Allanach for pointing this out.

One faces a fine-tuning problem, and needs additional assumptions to modify the measure appropriately.

We should point out that the only real input we use from the physical theory (standard model or a left-right symmetric extension of it) is, apart from the very definition of CP -violating parameters, how the theory restricts the type of mass matrices that appear. In particular, for any extension of the standard model (such as left-right symmetric models where parity is the left-right symmetry[‡]) that also has Hermitian mass matrices we would not see any modification in the results. Nevertheless, our calculations provide another example of how symmetries, in this case the presence or absence of a left-right symmetry, influence physical predictions of a given theory.

The structure of the remaining part of the chapter is as follows: In section 6.2, we review CP violation in the electroweak sector and detail why mass matrices may be assumed to be Hermitian in the standard model, but not if one has an extended left-right symmetry. In section 6.3, we compute expectation values for the Jarlskog invariant J using several of the measures on the space of CKM matrices that we have introduced. We also perform a more detailed analysis of how unlikely the observed value for J appears in these distributions. In section 6.4 we focus on the space of mass matrices instead, detailing how the observed values for the quark masses can be incorporated into a measure and how this completely changes predictions about likely values for J . We briefly mention how our analysis relates to the case of neutrinos and close with a summary and discussion of our findings.

6.2 CP Violation in the Standard Model and Beyond

We summarise how CP violation arises, first in the standard model and then in the more general case of left-right symmetric extensions, essentially following [91]. In the standard model, the quark fields appear as left-handed $SU(2)$ doublets and right-handed $SU(2)$ singlets:

$$\begin{pmatrix} q_{jL} \\ q'_{jL} \end{pmatrix}, \quad q_{jR}, \quad q'_{jR}, \quad j = 1, 2, \dots, N. \quad (6.2.1)$$

Here N is the number of quark families, which is arbitrary in the standard model, and normally taken to be three. The fields are written in a flavour basis which can be considered unphysical, since flavour eigenstates do not correspond to mass eigenstates.

[‡]We thank an anonymous referee for Phys. Rev. D for clarification on this point.

The coupling of the Higgs doublet H to quarks is given by

$$\mathcal{L}_{\text{Higgs}} = \sum_{j,k=1}^N \left(Y_{jk} \overline{(q, q')_{jL}} H^C q_{kR} + Y'_{jk} \overline{(q, q')_{jL}} H q'_{kR} + \text{h.c.} \right), \quad (6.2.2)$$

where the Higgs doublet H and its C -conjugate H^C can be written as

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_0 + i\phi_3 \end{pmatrix}, \quad H^C = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_0 - i\phi_3 \\ -\phi_1 + i\phi_2 \end{pmatrix}, \quad (6.2.3)$$

and Y_{jk} and Y'_{jk} are complex (Yukawa) couplings. H transforms under $SU(2)$ as a doublet and the term $\mathcal{L}_{\text{Higgs}}$ is invariant under $SU(2)$. This symmetry is then spontaneously broken by the Higgs potential $V(H)$ which gives rise to a vacuum expectation value v for ϕ_0 . The remaining components ϕ_j are ‘eaten’ by the W and Z bosons, which become massive, and the only remaining terms involve $\phi'_0 = \phi_0 - v$, giving masses to the quark fields:

$$\mathcal{L}_{\text{Higgs}} \xrightarrow{\text{SSB}} - \sum_{j,k=1}^N \left(m_{jk} \overline{q_{jL}} q_{kR} + m'_{jk} \overline{q'_{jL}} q'_{kR} + \text{h.c.} \right) \left(1 + \frac{1}{v} \phi'_0 \right). \quad (6.2.4)$$

The new **mass matrices** m and m' are (up to a factor of $\sqrt{2}$) given by the original Yukawa couplings, multiplied by v . Thus, it seems appropriate to regard either the set of Yukawa couplings together with the Higgs vacuum expectation value or the collection of elements of m and m' as fundamental parameters of the theory.

This part of the Lagrangian is formally CP invariant if and only if m and m' , which are so far arbitrary complex matrices, are real. Since this condition is not satisfied in Nature, one has formal CP violation. However, as remarked before, the Lagrangian has been written in the unphysical flavour basis. One can, for general m and m' , pass to a different basis, namely the basis of mass eigenstates, by diagonalising the mass matrices with unitary matrices:

$$m = U_L^\dagger \Delta U_R, \quad m' = (U'_L)^\dagger \Delta' U'_R, \quad (6.2.5)$$

thus the basis of mass eigenstates is related to the previously considered basis by

$$q_{L\text{phys}} = U_L q_L, \quad q_{R\text{phys}} = U_R q_R, \quad \text{etc.} \quad (6.2.6)$$

It is always possible to choose the unitary matrices so that Δ and Δ' are real, and in this new basis this part of the Lagrangian is invariant under C and P separately,

and hence also under CP . However, the electroweak Lagrangian also contains charged current terms mixing up- and down-type quarks, coupled to the W boson fields via (we are now using the basis of mass eigenstates)

$$X_C := (W_\mu^1 - iW_\mu^2)J_c^\mu + \text{h.c.}, \quad J_c^\mu := \overline{(u, c, t)_L} \gamma^\mu V \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix}, \quad (6.2.7)$$

where $V := U_L(U'_L)^\dagger$ is the **Cabibbo-Kobayashi-Maskawa** (CKM) matrix. In the basis of mass eigenstates, this term X_C is not invariant under CP unless V is real. Since we consider the mass eigenstates as physical, CP is violated through these charged current terms.

An important observation made in [55] is that one can redefine the right-handed quark bases by arbitrary unitary transformations,

$$U_R \rightarrow OU_R, \quad U'_R \rightarrow O'U'_R, \quad (6.2.8)$$

obtaining a new basis which is to be regarded as equally physical. This is due to the absence of charged current terms involving right-handed quarks, since they are singlets under $SU(2)$. It is therefore no loss of generality to set $U_R = U_L$ and $U'_R = U'_L$ in (6.2.5), and to assume that m and m' are Hermitian.

A natural way to extend the standard model is to assume the existence of a second $SU(2)$ symmetry which acts on the right-handed quarks, as in the Pati-Salam model [128]. In such extensions, one adds a term (6.2.7) for right-handed quarks to the Lagrangian. This has the important consequence that a general transformation (6.2.8) for arbitrary unitary transformations O and O' can no longer be regarded as giving an equivalent quark basis, since it modifies this new charged current term. The mass matrices cannot in general be taken to be Hermitian, but are arbitrary complex matrices.

There are now also two possibly CP violating terms, and two CKM matrices. We focus on $V = U_L(U'_L)^\dagger$ which involves the processes that are actually observed [89] and disregard $V_R = U_R(U'_R)^\dagger$ in the following.

Mathematically, V is an element of $SU(N)$, but since the phases of the quark fields are arbitrary (even for non-Hermitian mass matrices), V is only defined up to left or right multiplication by a diagonal element of $SU(N)$, *i.e.* an element of the maximal torus $U(1)^{N-1}$. The space of CKM matrices is therefore the double quotient $U(1)^{N-1} \backslash SU(N) / U(1)^{N-1}$, characterised by $(N-1)^2$ parameters, out of

which $\frac{1}{2}N(N-1)$ may be taken to be real (Euler) angles and the remaining $\frac{1}{2}(N-1)(N-2)$ appear as complex phases. It follows that the matrix V can be taken to be real for $N=2$, and so in the formalism explained here one needs at least three quark families to have a possibility of CP violation. Kobayashi and Maskawa were awarded the 2008 Nobel Prize for using this observation to predict the existence of a third quark family [98]. We set $N=3$ in what follows.

The mathematical theory of observable measures of CP violation in the standard model was developed by Jarlskog [91]. She showed [89] that all necessary and sufficient conditions for CP violation can be summarised as the following condition on the commutator of the Hermitian matrices m, m' :

$$\det C := \det(-i[m, m']) \neq 0. \quad (6.2.9)$$

One finds that explicitly

$$\det C = -2J(m_t - m_c)(m_c - m_u)(m_u - m_t)(m_b - m_s)(m_s - m_d)(m_d - m_b), \quad (6.2.10)$$

where $J := \Im(V_{11}V_{22}V_{12}^*V_{21}^*)$ is the **Jarlskog invariant** which is invariant under left or right multiplication of V by a diagonal matrix, *i.e.* an element of $U(1)^2$. The geometrical interpretation of the quantity J is given by the so-called unitarity triangles. These express the requirement on V to be unitary, so that for instance

$$(VV^\dagger)_{12} = V_{11}V_{21}^* + V_{12}V_{22}^* + V_{13}V_{23}^* = 0. \quad (6.2.11)$$

In the complex plane, the three complex numbers that sum to zero form the sides of a triangle. The absolute value $|J|$ is twice the area of this triangle. Since there are different unitarity triangles corresponding to different elements of VV^\dagger , all with the same area, there are several ways of expressing J in terms of the elements of V . A general formula is given by [89]

$$J \sum_{\gamma, l} \epsilon_{\alpha\beta\gamma} \epsilon_{jkl} = \Im(V_{\alpha j} V_{\beta k} V_{\alpha k}^* V_{\beta j}^*). \quad (6.2.12)$$

The quantities describing the CKM matrix which are invariant under rephasing of the quark fields are J and the absolute values $|V_{\alpha j}|$.

In the general case where m and m' are not assumed to be Hermitian, the corresponding quantity would be

$$\begin{aligned} \det \mathbf{C} &:= \det(-i[mm^\dagger, m'm'^\dagger]) \\ &= -2J(m_t^2 - m_c^2)(m_c^2 - m_u^2)(m_u^2 - m_t^2)(m_b^2 - m_s^2)(m_s^2 - m_d^2)(m_d^2 - m_b^2), \end{aligned} \quad (6.2.13)$$

which of course leads to the same conditions on the mass matrices as before (given that we took all masses to be positive before). In the literature, the use of \mathbf{C} is perhaps more common than the use of C , and one may well argue that this second measure of CP violation should be considered more fundamental as its value does not depend on the arbitrary signs of the mass terms in the Lagrangian.

Recall from chapter 5 the parametrisation of the CKM matrix (5.1.6)

$$V = \begin{pmatrix} \cos y \cos z & \cos y \sin z & e^{-iw} \sin y \\ -\cos x \sin z - e^{iw} \sin x \sin y \cos z & \cos x \cos z - e^{iw} \sin x \sin y \sin z & \sin x \cos y \\ \sin x \sin z - e^{iw} \cos x \sin y \cos z & -\sin x \cos z - e^{iw} \cos x \sin y \sin z & \cos x \cos y \end{pmatrix}, \quad (6.2.14)$$

where the ranges of the Euler angles x, y, z and the complex phase w are

$$0 \leq x, y, z \leq \frac{\pi}{2}, \quad 0 \leq w < 2\pi. \quad (6.2.15)$$

In this parametrisation, the Jarlskog invariant J is given by

$$J = \frac{1}{4} \sin(2x) \sin(2z) \sin y \cos^2 y \sin w. \quad (6.2.16)$$

It appears that the observed value for J is very small, since the maximal value would be $\frac{1}{6\sqrt{3}} \approx 0.1$, whereas in Nature [7]

$$J = 3.05_{-0.20}^{+0.19} \times 10^{-5}. \quad (6.2.17)$$

This discrepancy in 3.5 orders of magnitude is the fine-tuning problem we are investigating in this chapter.

In a general discussion where the values of the quark masses are not fixed, J is not an appropriate measure of CP violation, since even with nonvanishing J one could have CP conservation if, for example, the up and charm quark masses were coinciding. It was suggested in [90] to use an appropriately normalised form of $\det C$, namely

$$a_{CP} = 3\sqrt{6} \frac{\det C}{(\text{Tr } C^2)^{3/2}} \quad (6.2.18)$$

for three quark families as the unique basis independent measure of CP violation. This is a dimensionless number which takes values between -1 and $+1$, and is again observed to be very close to zero. When written out in terms of the CKM matrix parameters and quark masses, it is a rather complicated expression which is therefore not extremely useful in practical computations. In the present analysis, we assume the quark masses as known and regard J as the measure of CP violation.

6.3 Statistics on the Space of CKM Matrices

Let us compute expectation values for J for different choices of measure on the double quotient $U(1)^2 \backslash SU(3)/U(1)^2$, parametrised by coordinates (x, y, z, w) . Obviously, given a metric g , the expectation value of any function $f(x, y, z, w)$ is

$$\langle f \rangle = \frac{\int dx dy dz dw \sqrt{\det g} f(x, y, z, w)}{\int dx dy dz dw \sqrt{\det g}}. \quad (6.3.1)$$

As we mentioned in chapter 5, one possible viewpoint of a given function on $U(1)^2 \backslash SU(3)/U(1)^2$ is to see it as function on the homogeneous space $SU(3)/U(1)^2$ which is independent of the coordinates parametrising the left $U(1)^2$ (p and q in our conventions). Since we also saw that there is a unique left-invariant measure on the homogeneous space, this gives a unique definition of expectation values of such functions.

We computed this left-invariant measure on $SU(3)/U(1)^2$ to be

$$\sqrt{g} = N \sin(2x) \sin(2z) \sin y \cos^3 y \quad (6.3.2)$$

for some constant N .

Note that there is no need to pass from the double quotient to the homogeneous space if one regards the former as the fundamental domain of definition of a given function. One might equally well regard the left-invariance on the homogeneous space that has led to a unique measure as a spurious symmetry requirement.

Remembering that the CKM matrix is defined as $V = U(U')^\dagger$ in terms of two unitary matrices U and U' (denoted by U_L and U'_L for the general case of non-Hermitian mass matrices above), one might think about defining the Jarlskog invariant J in terms of coordinates on the Cartesian product of $SU(3) \times SU(3)$, the spaces of U and U' .[§] The natural measure on $SU(3) \times SU(3)$ is, in terms of left-invariant forms σ_i and σ'_i ,

$$\mu = N \sigma_1 \wedge \sigma_2 \wedge \sigma_3 \wedge \sigma_4 \wedge \sigma_5 \wedge \sigma_6 \wedge \sigma_7 \wedge \sigma_8 \wedge \sigma'_1 \wedge \sigma'_2 \wedge \sigma'_3 \wedge \sigma'_4 \wedge \sigma'_5 \wedge \sigma'_6 \wedge \sigma'_7 \wedge \sigma'_8. \quad (6.3.3)$$

Since it is only $V = UU'^\dagger$ that enters into the CP violating parameters, one could consider U and V as independent variables, *i.e.* write $U' = V^\dagger U$ for some matrix V .

[§] Ambiguities in the definition of U and U' mean that they should be regarded as living in quotient spaces such as $U(1)^2 \backslash SU(3)$, but since our left-invariant measures do not depend on the left and right $U(1)^2$ parameters such subtleties may be ignored.

Then the Maurer-Cartan form on the second $SU(3)$ is

$$i \lambda_a \sigma'_a \equiv U'^{\dagger} dU' = U^{\dagger} dU - U^{\dagger} (dV V^{\dagger}) U, \quad (6.3.4)$$

which gives $\sigma'_a = \sigma_a - h_{ab} \tau_b$, where τ_b are right-invariant forms on $SU(3)$ in terms of V coordinates and h_{ab} only depends on the U coordinates. The measure (6.3.3), expressed in terms of V and U coordinates, is thus a product of a function of the U coordinates and the natural measure in V coordinates (left- and right-invariant forms on $SU(3)$ give the same measure). Integration over the U coordinates then just gives an irrelevant constant, and one is left with the measure (6.3.2) on the space of V matrices. This justifies the use of (6.3.2) instead of the more complicated constructions obtained by reducing to the double quotient, and we will regard (6.3.2) as the most natural choice of measure on the parameter space.

For the measure (6.3.2) the evaluation of the necessary integrals is very simple and we find that all odd powers of J average to zero, and

$$\langle J^2 \rangle = \frac{1}{720} \approx 1.389 \times 10^{-3}, \quad \langle J^4 \rangle = \frac{1}{201600} \approx 4.960 \times 10^{-5}. \quad (6.3.5)$$

Thus we find that ΔJ for the Jarlskog invariant is given by

$$\Delta J = \frac{1}{12\sqrt{5}} \approx 0.0373, \quad (6.3.6)$$

which is about three orders of magnitude larger than the experimental value (6.2.17).

We argued in section 5.3 that a geometrically motivated measure on the double quotient space arises from Kaluza-Klein reduction of a left-invariant metric on $SU(3)$ or $SU(3)/U(1)^2$, and obtained the measure (5.3.7) from the bi-invariant metric on $SU(3)$. It is interesting how expectation values for powers of J compare to the previous calculations if this measure is used. Unfortunately, the expression for the measure is too complicated to allow us to perform the integrations analytically. Using numerical integration, we find that

$$\langle J^2 \rangle \approx 1.1161 \times 10^{-3}, \quad \langle J^4 \rangle \approx 3.750 \times 10^{-6}, \quad (6.3.7)$$

with the odd powers of J again averaging to zero. Thus we find

$$\Delta J \approx \sqrt{\langle J^2 \rangle} \approx 0.03341, \quad (6.3.8)$$

which is very close to the previous result.

Naively, one might have thought that since J is independent of all the $U(1)$ phases, the results would be the same whether one averaged over the space $U(1)^2 \backslash SU(3)/U(1)^2$, or else the flag manifold $SU(3)/U(1)^2$. Of course we know that this is not correct, since, as we have seen, the measure for the biquotient takes a different form than the measure one would obtain from just integrating over the left $U(1)^2$ parameters. Nevertheless, this does not seem to have a great impact on expectation values for J .

We briefly mentioned in chapter 5 that one could construct more general measures on the double quotient by Kaluza-Klein reduction of more general left-invariant metrics on $SU(3)/U(1)^2$. While this makes the expression for the measure even more complicated, the effect on expectation values of J seems somewhat limited, unless one takes extreme values for the “squashing parameters”. The general conclusion, namely that there appears to be a fine-tuning problem, is unchanged.[¶]

Given that we argued that the measure used by Ozsváth and Schücking (5.3.1) was not appropriate from a geometric viewpoint, we should compute expectation values for this measure also. We find that again $\langle J \rangle = 0$ and

$$\langle J^2 \rangle = 35 \times 2^{-16} \approx 5.341 \times 10^{-4}, \quad \langle J^4 \rangle = 27027 \times 2^{-34} \approx 1.573 \times 10^{-6}, \quad (6.3.9)$$

and that the standard deviation is

$$\Delta J^2 = \sqrt{\langle J^4 \rangle - \langle J^2 \rangle^2} = \sqrt{22127} \times 2^{-17} \approx 1.135 \times 10^{-3}, \quad (6.3.10)$$

and

$$\Delta J = \sqrt{\langle J^2 \rangle} = \frac{\sqrt{35}}{256} \approx 0.02311. \quad (6.3.11)$$

Again, the results are rather similar to the previous cases.

Assuming a uniform distribution over the angles, and hence treating the double coset as a flat four-dimensional manifold so that the measure is simply $\sqrt{\det g} = 1$, would give

$$\langle J \rangle = 0, \quad \langle J^2 \rangle = \frac{1}{2048} \approx 4.883 \times 10^{-4}, \quad \langle J^4 \rangle = 189 \times 2^{-27} \approx 1.408 \times 10^{-6}. \quad (6.3.12)$$

Hence, this simplest possible choice gives

$$\Delta J \approx 0.02210. \quad (6.3.13)$$

[¶]These calculations, mainly carried out by Chris Pope, are not included in detail in this thesis, but can be found in [67].

We have to conclude that, although we argued that some choices of measure on the space of CKM matrices are better motivated geometrically than others, if one is only interested in predictions about the magnitude of J there seems to be little difference between all of the choices we have examined. In all cases one would expect the observed value of J to be at least of order 10^{-2} , which leaves three orders of magnitude between prediction and observation.

6.3.1 Fine-tuning of J

In the previous subsection we saw that different measures on the space of mixing angles all seem to lead to expectation values for J which are about three orders of magnitude larger than the observed value. The value for J that we observe hence appears to be finely tuned. In this section we shall do a closer, mainly numerical, analysis of the fine-tuning involved. We compare results obtained by taking the $SU(3)$ -invariant (descending from the left-invariant measure on $SU(3)/U(1)^2$) and Kaluza-Klein (obtained from Kaluza-Klein reduction of the bi-invariant metric on $SU(3)$) measures, which seem natural from a geometric perspective, with a uniform distribution which is just the simplest possible choice.

The observed value for the Jarlskog invariant J is

$$J \approx 10^{-4.51} \approx e^{-10.39}. \quad (6.3.14)$$

In order to obtain a probability distribution for J we have used MATHEMATICA to numerically compute integrals of the form

$$\int dx dy dz dw \sqrt{\det g} \theta(a - |J|) \theta(|J| - b) \equiv P(b \leq |J| \leq a) \cdot \int dx dy dz dw \sqrt{\det g} \quad (6.3.15)$$

using Monte Carlo methods. The $SU(3)$ -invariant measure and the Kaluza-Klein measure disfavour small values of J more strongly than a uniform distribution would. For example, we obtain

$$P_{\text{flag}}(|J| \leq 10^{-4}) \approx 0.25\%, \quad P_{\text{KK}}(|J| \leq 10^{-4}) \approx 0.44\%. \quad (6.3.16)$$

Taking a uniform distribution $\sqrt{\det g} \equiv 1$, we get

$$P_{\text{unif}}(|J| \leq 10^{-4}) \approx 7\%. \quad (6.3.17)$$

In figures 6.1 to 6.3, we have plotted the probability distribution of $|J|$ logarithmically, together with a fit to a power law $p(|J|) \propto |J|^\lambda$ assumed to be valid for small $|J|$.

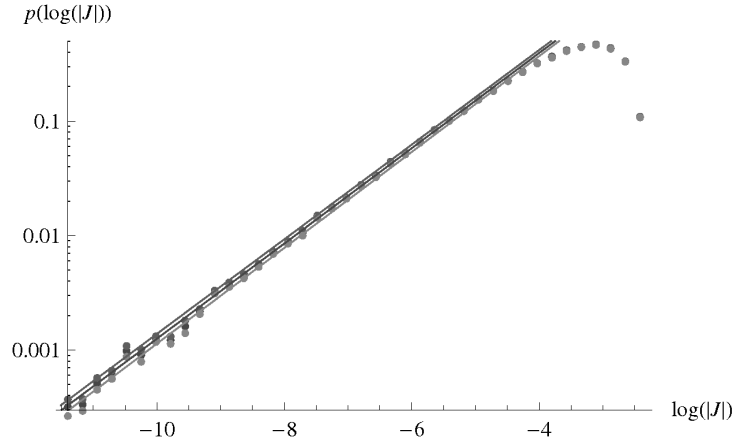


Figure 6.1: Probability distribution for $\log |J|$ using the $SU(3)$ -invariant flag measure.

The degree of fine-tuning required to reproduce a very small J is considerably higher if one uses the measure induced by a $SU(3)$ -invariant flag metric or the Kaluza-Klein metric, maybe contrary to what one might expect. Values of J close to its maximal value of $\frac{1}{6\sqrt{3}} \approx 0.0962$ are disfavoured in both cases. This motivated the use of a logarithmic scale for $|J|$.

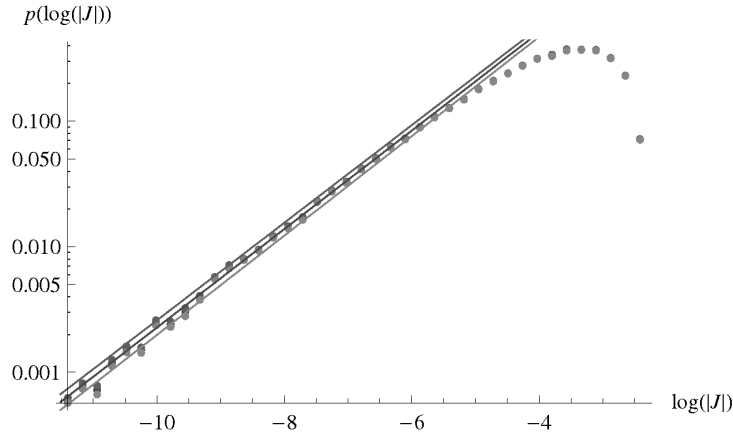


Figure 6.2: Probability distribution for $\log |J|$ using the Kaluza-Klein measure.

In all three cases the numerical results for small $|J|$ are well approximated by a power law of the form $p(|J|) = \alpha \cdot |J|^\lambda$ for the probability density of $|J|$. The logarithmic graphs show $p(\log |J|) \propto |J|^{\lambda+1}$. For the $SU(3)$ -invariant flag measure (Fig. 6.1), the best fit to the data in the region below $|J| = 10^{-2.3}$ or $\log |J| = -5.3$ is

$$\lambda_{\text{flag}} = -0.042(\pm 0.006), \quad \alpha_{\text{flag}} = 18.1(\pm 0.7); \quad (6.3.18)$$

for the Kaluza-Klein measure (Fig. 6.2) we fitted the data in the region below $|J| = 10^{-2.7}$ or $\log |J| = -6.2$ and obtained

$$\lambda_{\text{KK}} = -0.097(\pm 0.008), \quad \alpha_{\text{KK}} = 18.9(\pm 1.0); \quad (6.3.19)$$

finally for the uniform measure (Fig. 6.3), the best fit to the data in the region below $|J| = 10^{-3.4}$ or $\log |J| = -7.8$ is

$$\lambda_{\text{unif}} = -0.500(\pm 0.005), \quad \alpha_{\text{unif}} = 3.51(\pm 0.15). \quad (6.3.20)$$

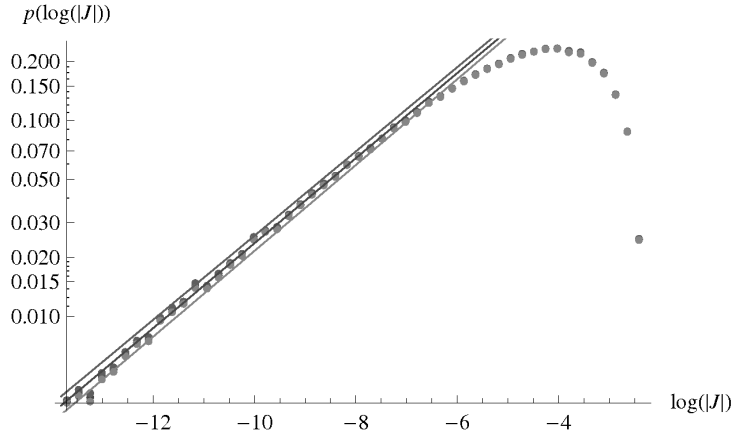


Figure 6.3: Probability distribution for $\log |J|$ using a uniform distribution.

6.3.2 Wolfenstein Parametrisation

A different parametrisation of the CKM matrix which is frequently used in the particle physics literature was introduced by Wolfenstein and is based on the experimentally observed hierarchy

$$y \ll x \ll z \ll 1 \quad (6.3.21)$$

in the mixing angles. One rewrites [169]

$$\sin z = \lambda, \quad \sin x = A\lambda^2, \quad \sin ye^{-i\omega} = A\lambda^3(\rho - i\eta) \quad (6.3.22)$$

and treats λ as a small parameter while A, ρ , and η are supposed to be parameters of order unity. In the modern literature one also frequently uses $\bar{\rho}, \bar{\eta}$ instead of ρ and

η because then the combination $\bar{\rho} + i\bar{\eta}$ is independent of the phase convention in the CKM matrix [7]. These parameters are defined by

$$\begin{aligned}\rho &= \sqrt{\frac{1 - A^2\lambda^4}{1 - \lambda^2}} \frac{\bar{\rho} - A^2\lambda^4(\bar{\rho}^2 + \bar{\eta}^2)}{(1 - A^2\lambda^4\bar{\rho})^2 + A^4\lambda^8\bar{\eta}^2}, \\ \eta &= \sqrt{\frac{1 - A^2\lambda^4}{1 - \lambda^2}} \frac{\bar{\eta}}{(1 - A^2\lambda^4\bar{\rho})^2 + A^4\lambda^8\bar{\eta}^2}.\end{aligned}\quad (6.3.23)$$

The experimental values for $\lambda, A, \bar{\rho}, \bar{\eta}$ are [7]^{||}

$$\lambda = 0.2272 \pm 0.0010, \quad A = 0.818^{+0.007}_{-0.017}, \quad \bar{\rho} = 0.221^{+0.064}_{-0.028}, \quad \bar{\eta} = 0.340^{+0.017}_{-0.045}. \quad (6.3.24)$$

One viewpoint on the Wolfenstein parametrisation is that it is adapted to the values for the CKM matrix entries that we observe and has no deeper significance; but often the viewpoint is expressed that this parametrisation expresses some kind of “natural hierarchy” in the mixing angles coming from physics beyond the standard model (see *e.g.* [24]). Treating the other parameters as “naturally of order unity” reduces our calculations to a one-dimensional problem as everything is only expanded in terms of λ . We find that the $SU(3)$ -invariant measure on the flag manifold is now, to leading order in λ ,

$$\left| \frac{\partial(x, y, z, w)}{\partial(\lambda, A, \bar{\rho}, \bar{\eta})} \right| \sqrt{g} \propto A^3 \lambda^{11} (1 + \lambda^2 + O(\lambda^4)), \quad (6.3.25)$$

and the Jarlskog invariant J is

$$\begin{aligned}J &= \frac{A^2 \bar{\eta} \lambda^6 (1 - A^2 \lambda^4) (1 - \lambda^2 - 2A^2 \bar{\rho} \lambda^4 - A^2 (\bar{\eta}^2 + (\bar{\rho} - 2)\bar{\rho}) \lambda^6 + A^4 (\bar{\eta}^2 + \bar{\rho}^2) \lambda^8)}{(1 - \lambda^2) (1 - 2A^2 \bar{\rho} \lambda^4 + A^4 (\bar{\eta}^2 + \bar{\rho}^2) \lambda^8)^2} \\ &= A^2 \bar{\eta} \lambda^6 + O(\lambda^{10}).\end{aligned}\quad (6.3.26)$$

Inverting this expression to leading order gives the probability distribution for J

$$p(J) \propto \frac{J}{A \bar{\eta}^2} \left(1 + \left(\frac{J}{A^2 \bar{\eta}} \right)^{1/3} + O(J^{2/3}) \right), \quad (6.3.27)$$

which is incompatible with the numerical results. Trying to improve this approximate result by letting $A, \bar{\rho}$ and $\bar{\eta}$ take all possible values leads to inconsistencies since the expansion in powers of J contains poles of arbitrary order in A . From our present viewpoint, where no mechanism for fixing these parameters close to one is known, the Wolfenstein parametrisation seems rather misleading when discussing geometric probability.

^{||}Note that only even powers of λ appear in all expansions, so that it is $\lambda^2 \approx 0.05$ which is the small parameter.

6.4 Statistics on the Space of Mass Matrices

In the previous sections we have focussed on $U(1)^2 \backslash SU(3)/U(1)^2$, the space of CKM matrices, as the space of CP violating parameters. Since $SU(3)$ is compact, this space has finite volume for a natural measure. But the CKM matrix is derived from the quark mass matrices, which could be viewed as more fundamental and more directly determined by physics beyond the standard model. In this section, we try to obtain statistics of the Jarlskog invariant J from a random distribution on the space of mass matrices; in the standard model, this is the space of 3×3 Hermitian matrices, whereas in certain extensions with left-right symmetry one has to take all 3×3 complex matrices into account.

In this section, as suggested in [89], we define a dimensionless mass matrix M by $M = m/\Lambda$, where m is the quark mass matrix in the notation of section 6.2, and Λ is a scale which may be chosen for convenience. A natural choice would be $\Lambda = m_t$ for the up-type or $\Lambda' = m_b$ for the down-type quarks, but since we in principle allow arbitrary values for the quark masses we leave Λ arbitrary.

We now need to pick a measure on the space of mass matrices in order to do statistics. Let us summarise the main results of section 5.4: For a given Hermitian matrix M , one defines a unitary matrix U and a real diagonal matrix D by

$$M_H = U^\dagger D U, \quad (6.4.1)$$

and the natural measure, invariant under conjugation by $U(3)$, on the space of such matrices was found to be

$$\begin{aligned} DM_H &:= (D_1 - D_2)^2 (D_1 - D_3)^2 (D_2 - D_3)^2 \sin(2x) \cos^3 y \sin y \sin(2z) \\ &\quad \times dD_1 dD_2 dD_3 dx dy dz dw dr dt. \end{aligned} \quad (6.4.2)$$

We noted that to avoid overcounting one should integrate over \mathbb{R}^3 times a quotient of $SU(3)$, more precisely $(U(1)^2 \times \mathfrak{S}_3) \backslash SU(3)$.

Let us contrast this with the space of all complex matrices; here we used the representation

$$M_c = U_L^\dagger D U_R, \quad (6.4.3)$$

and identified the relevant space one has to integrate over as $\mathbb{R}_+^3 \times (U(1)^3 \times \mathfrak{S}_3) \backslash (U(3) \times U(3))$. We then argued that if we are only interested in functions that depend on U_L , but not on U_R , such as the Jarlskog invariant J , we can integrate over the space of

U_R . Then the resulting measure on $\mathbb{R}_+^3 \times (U(1)^2 \times \mathfrak{S}_3) \backslash SU(3)$ is

$$DM_c := \prod_{i < j} (D_i^2 - D_j^2)^2 |D_1 D_2 D_3| \sin(2x) \cos^3 y \sin y \sin(2z) \\ \times dD_1 dD_2 dD_3 dx dy dz dw dr dt, \quad (6.4.4)$$

which obviously differs from (6.4.2) by different powers of the real numbers D_i that had the role of eigenvalues for Hermitian matrices. Since the measure (6.4.4) is invariant under $D_i \rightarrow -D_i$, we will integrate over all of \mathbb{R}^3 for simplicity. This only leads to an extra factor 8 which drops out in expectation values.

It is clear that in both cases the integrals over \mathbb{R}^3 with the given measures will diverge. We could introduce a cutoff for the quark masses, but then any expectation values for quark masses would strongly contradict observation, as there is no way to explain the observed mass hierarchy.

To both make the integrals converge and introduce a possibility for introducing the quark masses as additional parameters into our distribution, we introduce a function decaying sufficiently fast for large $|D_i|$ into the measure. A natural and simple choice is a Gaussian.

Since there are actually two integrations over the space of mass matrices, corresponding to two mass matrices M and M' for the up-type and down-type quarks, we will use the measure**

$$DM DM' \exp(-\text{Tr}(MM^\dagger A) - \text{Tr}(M'(M')^\dagger A')) \\ = DM DM' \exp(-\text{Tr}(U^\dagger D^2 U A) - \text{Tr}((U')^\dagger (D')^2 U' A')), \quad (6.4.5)$$

which for complex matrices is still invariant under the two right $U(3)$ actions on the complex matrices M and M' , but the invariance under the left $U(3)$ actions is broken to the subgroups commuting with A and A' respectively. We will assume A and A' to be Hermitian with non-negative eigenvalues, satisfying $[A, A'] = 0$ and use the redefinitions $U \rightarrow UW$, $U' \rightarrow U'W$, which leave J invariant, to diagonalise A and A' . Hence we can assume A and A' to be diagonal with positive entries in the following, and the subgroup commuting with A and A' is $U(1)^3$. For Hermitian matrices, (6.4.5) becomes

$$DM_H DM'_H \exp(-\text{Tr}(M_H^2 A)) \exp(-\text{Tr}((M'_H)^2 A)), \quad (6.4.6)$$

** DM is to be replaced by DM_H and DM_c for Hermitian and complex matrices respectively, and from now on we write U instead of U_L in the case of complex matrices as well, since U_R has disappeared from our calculations.

where the Gaussian breaks the symmetry of the measure from invariance under $(U(3) \times U(3))$ to invariance under $(U(1)^3 \times U(1)^3)$, again the subgroup commuting with A and A' .

We will find that this symmetry breaking is necessary to obtain a distribution that can reproduce the observed quark masses. Our proposal is to fit the diagonal matrices A and A' to the observed quark masses and to use the resulting probability distribution for statistics of J .

An expectation value of a quantity such as J^2 is then given by^{††}

$$\begin{aligned} \langle J^2 \rangle &= N \int DM DM' e^{-\text{Tr}(MM^\dagger A) - \text{Tr}(M' M'^\dagger A')} J^2(M, M') \\ &= N \int_{\mathbb{R}^6} dD dD' \int_{((U(1)^2 \times \mathfrak{S}_3) \backslash SU(3))^2} DU DU' e^{-\text{Tr}(D^2 U A U^\dagger) - \text{Tr}((D')^2 U' A' U'^\dagger)} J^2. \end{aligned} \quad (6.4.7)$$

Here DU and DU' are the measures on $(U(1)^2 \times \mathfrak{S}_3) \backslash SU(3)$, given by (5.1.12),

$$DU := \sin(2x) \sin y \cos^3 y \sin(2z), \quad (6.4.8)$$

and

$$dD := \begin{cases} (D_1 - D_2)^2 (D_1 - D_3)^2 (D_2 - D_3)^2 dD_1 dD_2 dD_3, & M \text{ Hermitian}, \\ (D_1^2 - D_2^2)^2 (D_1^2 - D_3^2)^2 (D_2^2 - D_3^2)^2 |D_1 D_2 D_3| dD_1 dD_2 dD_3, & M \text{ complex}, \end{cases} \quad (6.4.9)$$

and the normalisation factor N is defined by

$$\frac{1}{N} := \int_{\mathbb{R}^6} dD dD' \int_{((U(1)^2 \times \mathfrak{S}_3) \backslash SU(3))^2} DU DU' e^{-\text{Tr}(D^2 U A U^\dagger) - \text{Tr}((D')^2 U' A' U'^\dagger)}. \quad (6.4.10)$$

From $J = \Im(V_{11} V_{22} V_{12}^* V_{21}^*)$ and $V = U U'^\dagger$, we have

$$J(U, U') = \sum_{a,b,c,d=1}^3 \Im(U_{1a} U_{2b} U_{1c}^* U_{2d}^* U'_{1a} U'_{2b}^* U'_{2c} U'_{1d}). \quad (6.4.11)$$

At this point, it is perhaps instructive to note that choosing A proportional to the identity would split the integral (6.4.7) into a product of an integral over \mathbb{R}^6 , *i.e.* the quark masses, which just gives a constant, and an integral of J^2 over $((U(1)^2 \times \mathfrak{S}_3) \backslash SU(3))^2$. Since all even powers of J are invariant under the \mathfrak{S}_3 action on U and U' , this can be replaced by an integral over $(U(1)^2 \backslash SU(3))^2$ if averages are concerned. By the arguments presented in section 6.3, a change of coordinates reduces this to a

^{††}In all cases considered, all odd powers of J again have expectation value zero.

single integration over a copy of $SU(3)/U(1)^2$, and one recovers the results of section 6.3 for expectation values of powers of J . The introduction of more general diagonal matrices A and A' means that the maximal invariance of the measure $DM DM'$ is broken down to the action of a subgroup.

It should be clear from (6.4.7) that multiplying A (or A') by a constant is the same as rescaling the eigenvalues D_i (or D'_i) and so amounts to a rescaling of Λ (or Λ'). We can therefore, without any loss of generality, choose

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/\mu_c^2 & 0 \\ 0 & 0 & 1/\mu_u^2 \end{pmatrix}, \quad A' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/\mu_s^2 & 0 \\ 0 & 0 & 1/\mu_d^2 \end{pmatrix}, \quad (6.4.12)$$

where μ_c, μ_u, μ_s , and μ_d are dimensionless parameters that we are free to choose so as to reproduce the observed quark masses as expectation values. (In the case of an exponential $\exp(-\text{Tr}(D^2 A))$, these would of course be equal to the respective quark masses, expressed in units where $\Lambda = m_t$ and $\Lambda' = m_b$.) Because of experimental uncertainties in the up and quark masses, one can modify this distribution to reproduce different values for these masses.

6.4.1 Hermitian Mass Matrices

We shall now present the result of evaluating the integral (6.4.7), first for Hermitian and then for complex mass matrices, and therefore for the two different choices of measure (6.4.9). We will have to use various approximations which we try to justify in the following, and the final result is obtained independently by numerical integration and by an analytical approximation, where we find relatively good agreement between the two.

The need for approximations in the integral (6.4.7) is immediate since the expression for J in terms of coordinates on $((U(1)^2 \times \mathfrak{S}_3) \backslash SU(3))^2$ is too complicated to be given explicitly. However, since

$$\text{Tr}(D^2 U A U^\dagger) = \sum_a D_a^2 \sum_c A_c |U_{ac}|^2 =: \sum_a D_a^2 \xi_a, \quad \text{Tr}((D')^2 U' A' U'^\dagger) =: \sum_a (D'_a)^2 \xi'_a \quad (6.4.13)$$

with

$$\begin{aligned} \xi_1 &= A_1 \cos^2 y \cos^2 z + A_2 \cos^2 y \sin^2 z + A_3 \sin^2 y, \\ \xi'_1 &= A'_1 \cos^2 y' \cos^2 z' + A'_2 \cos^2 y' \sin^2 z' + A'_3 \sin^2 y', \end{aligned} \quad (6.4.14)$$

and we assume $A_3 \gg 1$ and $A'_3 \gg 1$, the integrand is negligibly small unless $y \approx 0$ and $y' \approx 0$. We use this to approximate the integrals over y and y' (recall the restriction of the coordinates on $U(1)^2 \backslash SU(3)$ given by (5.4.9) which restricts from $U(1)^2 \backslash SU(3)$ to $(U(1)^2 \times \mathfrak{S}_3) \backslash SU(3)$):

$$\begin{aligned}
& \int_0^{\arctan(\sin x)} dy \int_0^{\arctan(\sin x')} dy' \cos^3 y \sin y \cos^3 y' \sin y' e^{-\text{Tr}(D^2 U A U^\dagger) - \text{Tr}((D')^2 U' A' U'^\dagger)} J^2(U, U') \\
& \approx \int_0^\infty dy \int_0^\infty dy' y y' e^{-A_3 y^2 - A'_3 (y')^2} \cdot \left(e^{-\text{Tr}(D^2 U A U^\dagger) - \text{Tr}((D')^2 U' A' U'^\dagger)} J^2(U, U') \right) \Big|_{y=y'=0} \\
& \approx \frac{1}{4A_3 A'_3} \left(e^{-\text{Tr}(D^2 U A U^\dagger) - \text{Tr}((D')^2 U' A' U'^\dagger)} J^2(U, U') \right) \Big|_{y=y'=0}. \tag{6.4.15}
\end{aligned}$$

It turns out that this is independent of w and w' . Constant prefactors such as $1/4A_3 A'_3$ appearing in both numerator and denominator can be dropped, and so we have

$$\langle J^2 \rangle \approx \frac{\int_{\mathbb{R}^6} dD dD' \int d^4 x d^4 x' s_{2x} s_{2z} s_{2x'} s_{2z'} \left(e^{-\text{Tr}(D^2 U A U^\dagger) - \text{Tr}((D')^2 U' A' U'^\dagger)} J^2(U, U') \right) \Big|_{y=y'=0}}{\int_{\mathbb{R}^6} dD dD' \int d^4 x d^4 x' s_{2x} s_{2z} s_{2x'} s_{2z'} \left(e^{-\text{Tr}(D^2 U A U^\dagger) - \text{Tr}((D')^2 U' A' U'^\dagger)} \right) \Big|_{y=y'=0}}, \tag{6.4.16}$$

where $s_{2x} = \sin 2x$ etc. and

$$\int d^4 x \equiv \int_0^{\pi/4} dx \int_0^{\pi/2} dz \int_0^{2\pi} dr \int_0^{2\pi} dt \tag{6.4.17}$$

and similarly for $\int d^4 x'$.

Now we can integrate over both copies of \mathbb{R}^3 in (6.4.16), using

$$\begin{aligned}
& \int_{-\infty}^\infty dD_1 \int_{-\infty}^\infty dD_2 \int_{-\infty}^\infty dD_3 (D_1 - D_2)^2 (D_1 - D_3)^2 (D_2 - D_3)^2 e^{-\xi_1 D_1^2 - \xi_2 D_2^2 - \xi_3 D_3^2} \\
& = \frac{3\pi^{3/2}}{8\xi_1^{5/2} \xi_2^{5/2} \xi_3^{5/2}} (\xi_1^2 (\xi_2 + \xi_3) + \xi_2^2 (\xi_1 + \xi_3) + \xi_3^2 (\xi_2 + \xi_1) - 2\xi_1 \xi_2 \xi_3). \tag{6.4.18}
\end{aligned}$$

The explicit expression for J at $y = y' = 0$ is

$$\begin{aligned}
J(U, U') \Big|_{y=y'=0} & = \frac{1}{4} s_{2x} s_{2x'} \left\{ c_z^2 s_{z'}^3 \sin(3\hat{r} + \hat{t}) + c_z^3 c_{z'} s_z^2 \sin(3\hat{r} - \hat{t}) \right. \\
& \quad - c_z^2 s_z s_{z'} (c_z^2 [\sin(3\hat{r} + \hat{t}) + \sin(3\hat{r} - 3\hat{t})] - s_z^2 \sin(3\hat{r} + \hat{t})) \\
& \quad \left. + c_z c_{z'} s_z^2 (c_z^2 \sin(3\hat{r} - \hat{t}) - s_z^2 [\sin(3\hat{r} + 3\hat{t}) + \sin(3\hat{r} - \hat{t})]) \right\} \tag{6.4.19}
\end{aligned}$$

where $s_x = \sin x$, $c_{z'} = \cos z'$, etc., $\hat{r} = r - r'$, and $\hat{t} = t - t'$. Integrating (6.4.19) over r, r', t , and t' indeed gives zero, which is why we choose to use J^2 .

We need to determine the parameters appearing in the matrices A and A' in (6.4.12) by fitting them to the expectation values for quark masses that we want to reproduce. We first observe that expectation values for squared mass matrices take the relatively simple form

$$\langle D_1^2 \rangle \approx \frac{\int_{\mathbb{R}^3} dD D_1^2 \int_0^{\pi/4} dx \int_0^{\pi/2} dz \sin 2x \sin 2z \left(e^{-\text{Tr}(D^2 U A U^\dagger)} \right) \big|_{y=0}}{\int_{\mathbb{R}^3} dD \int_0^{\pi/4} dx \int_0^{\pi/2} dz \sin 2x \sin 2z \left(e^{-\text{Tr}(D^2 U A U^\dagger)} \right) \big|_{y=0}}. \quad (6.4.20)$$

The denominator of (6.4.20) is explicitly

$$I_D := \int_0^{\pi/4} dx \int_0^{\pi/2} dz \sin 2x \sin 2z \frac{3\pi^{3/2}}{8\xi_1^{5/2}\xi_2^{5/2}\xi_3^{5/2}} \times \\ (\xi_1^2(\xi_2 + \xi_3) + \xi_2^2(\xi_1 + \xi_3) + \xi_3^2(\xi_2 + \xi_1) - 2\xi_1\xi_2\xi_3), \quad (6.4.21)$$

where within our approximation the quantities ξ_a are taken at $y = 0$,

$$\xi_1 = A_1 \cos^2 z + A_2 \sin^2 z, \quad \xi_2 = A_1 \cos^2 x \sin^2 z + A_2 \cos^2 x \cos^2 z + A_3 \sin^2 x, \\ \xi_3 = A_1 \sin^2 x \sin^2 z + A_2 \sin^2 x \cos^2 z + A_3 \cos^2 x, \quad (6.4.22)$$

with $A_3 \gg A_2 \gg A_1$. We notice that all ξ_a are nonzero for all values of x and z . Furthermore, the integral is dominated by very small x and z (we cannot have $x = \frac{\pi}{2}$), and we can approximate I_D well by only keeping the terms of leading order in x and z in the trigonometric functions, and

$$\xi_1^2(\xi_2 + \xi_3) + \xi_2^2(\xi_1 + \xi_3) + \xi_3^2(\xi_2 + \xi_1) - 2\xi_1\xi_2\xi_3 \approx A_3^3 x^2 + A_3^2 A_2, \quad (6.4.23)$$

which are the leading terms (as we shall see, the first of these is effectively also of

order $A_3^2 A_2$):

$$\begin{aligned}
I_D &\approx \frac{3\pi^{3/2}}{8} \int_0^{\pi/4} dx \int_0^{\pi/2} dz 4xz (A_3^3 x^2 + A_3^2 A_2)(A_1 + A_2 z^2)^{-5/2} (A_2 + A_3 x^2)^{-5/2} A_3^{-5/2} \\
&\approx \frac{3\pi^{3/2}}{8} \int_0^\infty dX \int_0^\infty dZ (A_3^3 X + A_3^2 A_2)(A_1 + A_2 Z)^{-5/2} (A_2 + A_3 X)^{-5/2} A_3^{-5/2} \\
&= \frac{3\pi^{3/2}}{8} \int_0^\infty dX (A_3^3 X + A_3^2 A_2)(A_2 + A_3 X)^{-5/2} A_3^{-5/2} \cdot \frac{2}{3A_1^{3/2} A_2} \\
&= \frac{3\pi^{3/2}}{8} \frac{2}{3A_1^{3/2} A_2} \left(\frac{2}{3} A_2^{-1/2} A_3^{-3/2} + \int_0^\infty dX \frac{2}{3A_3^{1/2}} (A_2 + A_3 X)^{-3/2} \right) \\
&= \frac{\pi^{3/2}}{4A_1^{3/2} A_2} \left(\frac{2}{3} A_2^{-1/2} A_3^{-3/2} + \frac{4}{3A_3^{3/2} A_2^{1/2}} \right) = \frac{\pi^{3/2}}{2A_1^{3/2} A_2^{3/2} A_3^{3/2}}. \tag{6.4.24}
\end{aligned}$$

Similarly, we find for the numerator of (6.4.20)

$$\begin{aligned}
I_D \langle D_1^2 \rangle &\approx \frac{15\pi^{3/2}}{16} \int_0^\infty dX \int_0^\infty dZ (A_3^3 X + A_3^2 A_2)(A_1 + A_2 Z)^{-7/2} (A_2 + A_3 X)^{-5/2} A_3^{-5/2} \\
&= \frac{15\pi^{3/2}}{16} \int_0^\infty dX (A_3^3 X + A_3^2 A_2)(A_2 + A_3 X)^{-5/2} A_3^{-5/2} \cdot \frac{2}{5A_2 A_1^{5/2}} \\
&= \frac{3\pi^{3/2}}{4A_1^{5/2} A_2^{3/2} A_3^{3/2}}, \tag{6.4.25}
\end{aligned}$$

hence

$$\langle D_1^2 \rangle \approx \frac{3}{2A_1}. \tag{6.4.26}$$

Redoing the same calculation for D_2 and D_3 gives

$$\langle D_2^2 \rangle \approx \frac{1}{2A_2}, \quad \langle D_3^2 \rangle \approx \frac{1}{2A_3}. \tag{6.4.27}$$

There is a relative factor of 3 which has to be taken into account when determining A and A' .

Because of the dependence of masses on the energy scale in quantum field theory, described by the renormalisation group, there is some ambiguity in what is meant by the “quark masses” we want to reproduce. Following [144], for example, we take all the quark masses evolved to the scale of the Z boson mass. These are given in [170]:

$$\begin{aligned}
(m_u, m_c, m_t) &= (1.27_{-0.42}^{+0.50} \text{ MeV}, 0.619 \pm 0.084 \text{ GeV}, 171.7 \pm 3.0 \text{ GeV}); \\
(m_d, m_s, m_b) &= (2.90_{-1.19}^{+1.24} \text{ MeV}, 55_{-15}^{+16} \text{ MeV}, 2.89 \pm 0.09 \text{ GeV}). \tag{6.4.28}
\end{aligned}$$

We use the central values

$$\begin{aligned}(m_u, m_c, m_t) &:= (1.27 \text{ MeV}, 0.619 \text{ GeV}, 171.7 \text{ GeV}); \\ (m_d, m_s, m_b) &:= (2.9 \text{ MeV}, 55 \text{ MeV}, 2.89 \text{ GeV}).\end{aligned}\quad (6.4.29)$$

The mass scales Λ and Λ' that were so far arbitrary are fixed by setting $\langle D_1^2 \rangle = (m_t/\Lambda)^2$ and $\langle (D'_1)^2 \rangle = (m_b/\Lambda')^2$. By comparing the results obtained by numerical integration with the values we want to reproduce, we can then fix the parameters μ_c, μ_u, μ_s and μ_d .

In the case of the positively charged top, charm and up quarks, which exhibit a more extreme quark mass hierarchy, we find that numerical calculations (using MATHEMATICA) reproduce the results we have obtained analytically very well (see Table 6.1). For the negatively charged quarks, we find numerically that we have to use relative factors different from 3 to reproduce the observed masses. Comparing the numerical results with (6.4.29), we fix the parameters appearing in A and A' to

$$\begin{aligned}\mu_c^2 &= 3 \left(\frac{m_c}{m_t} \right)^2 \approx 3.90 \times 10^{-5}, \quad \mu_u^2 = 3 \left(\frac{m_u}{m_t} \right)^2 \approx 1.64 \times 10^{-10}, \\ \mu_s^2 &= \frac{3}{2} \left(\frac{m_s}{m_b} \right)^2 \approx 5.43 \times 10^{-4}, \quad \mu_d^2 = \frac{12}{5} \left(\frac{m_d}{m_b} \right)^2 \approx 2.42 \times 10^{-6}.\end{aligned}\quad (6.4.30)$$

As a brief side remark, we see that the dominant contributions to these integrals come from the regions

$$y \approx \sqrt{\frac{1}{A_3}}, \quad y' \approx \sqrt{\frac{1}{A'_3}}, \quad x \approx \sqrt{\frac{A_2}{A_3}}, \quad z \approx \sqrt{\frac{A_1}{A_2}}, \quad x' \approx \sqrt{\frac{A'_2}{A'_3}}, \quad z' \approx \sqrt{\frac{A'_1}{A'_2}}, \quad (6.4.31)$$

and these values are all small compared to one. This observation allows us to give rough estimates for magnitudes of individual elements of the CKM matrix.

In the standard convention the ordering of the quark families is (u, c, t) and not (t, c, u) as used in (6.4.12), which means that in our parametrisation,

$$|(UU'^\dagger)_{13}| = |V_{td}|, \quad |(UU'^\dagger)_{12}| = |V_{ts}|, \quad |(UU'^\dagger)_{23}| = |V_{cd}|. \quad (6.4.32)$$

Since all of the numbers in (6.4.31) are small, we only keep leading terms in the angles on U and U' :

$$\begin{aligned}|(UU'^\dagger)_{13}| &\approx |x'(z' - z) - e^{iw'}y' + \dots| \approx x'z' \approx \mu_d \approx 2 \times 10^{-3}, \\ |(UU'^\dagger)_{12}| &\approx z' \approx \mu_s \approx 0.02, \quad |(UU'^\dagger)_{23}| \approx x' \approx \frac{\mu_d}{\mu_s} \approx 0.07.\end{aligned}\quad (6.4.33)$$

Experimental values are [7]

$$|V_{td}| = (8.14_{-0.64}^{+0.32}) \times 10^{-3}, \quad |V_{ts}| = (41.61_{-0.78}^{+0.12}) \times 10^{-3}, \quad |V_{cd}| = 0.2271_{-0.0010}^{+0.0010}. \quad (6.4.34)$$

Our rough estimates reproduce the right ordering of the three parameters and are accurate to factors of order a few. A more careful analysis would involve computing expectation values for these parameters in the distribution we have assumed.

We return to the task of computing the expectation value of J^2 . In order to obtain an analytical expression, we use the fact that the main contribution to the integral (6.4.16) will come from small z to only take the term in (6.4.19) that is nonzero at $z = 0$. Averaging over $r, t, r',$ and t' gives a factor of $1/2$, as one might have expected, and therefore we use

$$J_{\text{small } z}^2 := \frac{1}{2} \sin^2 x \cos^2 x \sin^2 x' \cos^2 x' \cos^2 z' \sin^4 z' \quad (6.4.35)$$

for our calculations. The numerator of (6.4.16) is the product (using again that only small z contributes)

$$\begin{aligned} & \frac{9\pi^3}{32} \times \int_0^{\pi/2} dz \frac{2z}{(A_1 + A_2 z^2)^{5/2}} \times \int_0^{\pi/2} dz' \frac{\sin 2z' \cos^2 z' \sin^4 z'}{(A_1 \cos^2 z' + A_2 \sin^2 z')^{5/2}} \\ & \times \int_0^{\pi/4} dx \frac{\sin 2x \sin^2 x \cos^2 x (A_3^3 \cos^2 x \sin^2 x + A_3^2 A_2 (\cos^6 x + 2 \cos^2 x \sin^2 x + \sin^6 x))}{(A_2 \cos^2 x + A_3 \sin^2 x)^{5/2} (A_3 \cos^2 x + A_2 \sin^2 x)^{5/2}} \end{aligned} \quad (6.4.36)$$

times an integral identical to the second line of (6.4.36), except that A_2 and A_3 are replaced by A'_2 and A'_3 . The first two factors are $\frac{2}{3A_1^{3/2}A_2}$ and $\frac{4}{3\sqrt{A'_2}(\sqrt{A'_1} + \sqrt{A'_2})^4}$, respectively; for the second line of (6.4.36) we change variables to $X = \cos^2 x$ to obtain

$$\int_{1/2}^1 dX \frac{X(1-X)(A_3^3 X(1-X) + A_3^2 A_2 (X^2 - X + 1))}{(A_2 X + A_3 (1-X))^{5/2} (A_3 X + A_2 (1-X))^{5/2}} \approx \frac{1}{A_3^2} \left(\arctan \left(\frac{1}{2} \sqrt{\frac{A_3}{A_2}} \right) - 2 \sqrt{\frac{A_2}{A_3}} \right), \quad (6.4.37)$$

where we are dropping corrections of order $\frac{A_2}{A_3}$. Putting everything together, we obtain

$$\begin{aligned}
\langle J_{\text{small } z}^2 \rangle &\approx \frac{(A'_1)^{3/2} A'_2 \sqrt{A_2}}{\sqrt{A_3 A'_3} (\sqrt{A'_1} + \sqrt{A'_2})^4} \left(\arctan \sqrt{\frac{A_3}{4A_2}} - \sqrt{\frac{4A_2}{A_3}} \right) \\
&\times \left(\arctan \sqrt{\frac{A'_3}{4A'_2}} - \sqrt{\frac{4A'_2}{A'_3}} \right) \\
&= \frac{\frac{4}{\sqrt{15}} m_u m_d m_b}{m_c m_s^2 \left(1 + \sqrt{\frac{2}{3}} \frac{m_b}{m_s} \right)^4} \left(\arctan \frac{m_c}{2m_u} - \frac{2m_u}{m_c} \right) \\
&\times \left(\arctan \sqrt{\frac{5}{32}} \frac{m_s}{m_d} - \sqrt{\frac{32}{5}} \frac{m_d}{m_s} \right), \tag{6.4.38}
\end{aligned}$$

where the numerical factors appearing in the last line come from the different factors chosen in (6.4.30). Note that the top quark mass does not appear in this approximate result.

We also use numerical integration to check the validity of our analytical approximations. In these calculations we use both the simplified expression $J_{\text{small } z}^2$ and the expression for J given in (6.4.19). We find that for the first quantity, the numerically evaluated expectation value $\langle J_{\text{small } z}^2 \rangle$ is about 7/6 of (6.4.38), and the numerical result for $\langle J^2 \rangle$ (taken at $y = y' = 0$) is

$$\langle J^2 \rangle \approx 5.28 \times 10^{-9}, \tag{6.4.39}$$

which gives

$$\Delta J = \sqrt{\langle J^2 \rangle} \approx 7.27 \times 10^{-5} \tag{6.4.40}$$

which is much closer to the observed value than any of the previously obtained results. Assuming a Gaussian distribution for J which is peaked at zero, the probability of finding a small J , in the sense of Sec. 6.3.1, is now

$$P_{\text{mass}}(|J| \leq 10^{-4}) \approx 83\%, \tag{6.4.41}$$

whereas the probability of finding a J which is even smaller than the observed value is

$$P_{\text{mass}}(|J| \leq 3 \times 10^{-5}) \approx 32\%. \tag{6.4.42}$$

The observed value for J can no longer be viewed as being finely tuned if the distribution used in our calculations is assumed.

Table 6.1: Analytical and numerical results for integrals of interest.

Quantity	I_D (over x, z)	$I_D \langle D_1^2 \rangle$	$I_D \langle D_2^2 \rangle$	$I_D \langle D_3^2 \rangle$
Analytical result	1.43×10^{-21}	2.14×10^{-21}	2.78×10^{-26}	1.17×10^{-31}
Numerical result	1.43×10^{-21}	2.14×10^{-21}	2.78×10^{-26}	1.18×10^{-31}
Quantity	I'_D (over x', z')	$I'_D \langle (D'_1)^2 \rangle$	$I'_D \langle (D'_2)^2 \rangle$	$I'_D \langle (D'_3)^2 \rangle$
Analytical result	1.32×10^{-13}	1.99×10^{-13}	3.60×10^{-17}	1.60×10^{-19}
Numerical result	1.32×10^{-13}	1.98×10^{-13}	7.28×10^{-17}	1.99×10^{-19}
Quantity	\tilde{I}_D (over x, z, x', z')	$\langle J_{\text{small } z}^2 \rangle$	$\langle J^2 \rangle$	
Analytical result	1.89×10^{-34}	3.22×10^{-9}	—	
Numerical result	1.88×10^{-34}	3.73×10^{-9}	5.28×10^{-9}	

To test the sensitivity of our results to changes in the parameters, we take values at the upper or lower limit in (6.4.28) and try to find the highest and lowest values for $\langle J^2 \rangle$. We find that setting

$$\begin{aligned}
 (m_u, m_c, m_t) &:= (0.85 \text{ MeV}, 0.535 \text{ GeV}, 174.7 \text{ GeV}); \\
 (m_d, m_s, m_b) &:= (1.71 \text{ MeV}, 40 \text{ MeV}, 2.98 \text{ GeV})
 \end{aligned} \tag{6.4.43}$$

gives

$$\langle J^2 \rangle \approx 1.86 \times 10^{-9} \tag{6.4.44}$$

and

$$\Delta J = \sqrt{\langle J^2 \rangle} \approx 4.31 \times 10^{-5}, \tag{6.4.45}$$

whereas setting

$$\begin{aligned}
 (m_u, m_c, m_t) &:= (1.77 \text{ MeV}, 0.535 \text{ GeV}, 168.7 \text{ GeV}); \\
 (m_d, m_s, m_b) &:= (4.14 \text{ MeV}, 71 \text{ MeV}, 2.8 \text{ GeV})
 \end{aligned} \tag{6.4.46}$$

gives

$$\langle J^2 \rangle \approx 1.52 \times 10^{-8} \tag{6.4.47}$$

and

$$\Delta J = \sqrt{\langle J^2 \rangle} \approx 1.23 \times 10^{-4}. \tag{6.4.48}$$

While the result for ΔJ depends on the choice of parameters, we see that the general conclusion, which is a comparison of orders of magnitude, is unchanged by reasonable

modifications of the parameters (*i.e.* values for the quark masses). Even the greatest possible value for ΔJ is significantly lower than any of the values obtained in previous sections.

In this subsection, we have established that assuming the observed hierarchy in quark masses in a Gaussian distribution over the space of mass matrices gives expectation values for J^2 which are small enough to regard the observed value as “natural” and not finely tuned. This statistical observation seems to open up the possibility that the same mechanism that is responsible for the apparently unlikely hierarchy in quark masses might also explain why the observed value for $|J|$ is so small.

A more detailed analysis including a probability density for $|J|$ for this distribution is left to future work, since the numerical methods used here do not give sufficiently accurate results.

6.4.2 Complex Mass Matrices

The corresponding calculation for general complex mass matrices differs from the one for Hermitian mass matrices detailed in the previous subsection by a different measure over \mathbb{R}^6 , the space of possible quark masses. Hence after the integration over \mathbb{R}^6 has been performed, the resulting integrands will be different functions of the quantities ξ_a introduced in the previous section. The steps of calculating an expectation value for J^2 will be the same as in the previous subsection, and we will use similar analytical approximations as well. It suffices therefore to highlight those steps where the resulting formulae differ from the previous ones.

In the previous subsection we concluded that a natural measure on the space of Hermitian mass matrices that involves the observed quark masses reproduces the observed magnitude of CP violation very well. The quark mass hierarchy had the effect of significantly lowering the expectation value of J^2 . Looking at the different measures (6.4.9), we would expect this effect much stronger if general complex mass matrices are considered. This expectation will be confirmed: There now appears to be fine-tuning in the opposite way, *i.e.* one would expect CP violation to be much smaller than is observed.

Using the same approximation of small y and y' as previously, we are again left

with computing an integral of the form

$$\langle J^2 \rangle \approx \frac{\int_{\mathbb{R}^6} dD dD' \int d^4x d^4x' s_{2x} s_{2z} s_{2x'} s_{2z'} \left(e^{-\text{Tr}(D^2 U A U^\dagger) - \text{Tr}((D')^2 U' A' U'^\dagger)} J^2(U, U') \right) \Big|_{y=y'=0}}{\int_{\mathbb{R}^6} dD dD' \int d^4x d^4x' s_{2x} s_{2z} s_{2x'} s_{2z'} \left(e^{-\text{Tr}(D^2 U A U^\dagger) - \text{Tr}((D')^2 U' A' U'^\dagger)} \right) \Big|_{y=y'=0}}, \quad (6.4.49)$$

where

$$\int d^4x \equiv \int_0^{\pi/4} dx \int_0^{\pi/2} dz \int_0^{2\pi} dr \int_0^{2\pi} dt \quad (6.4.50)$$

and similarly for $\int d^4x'$. The symbol dD of course now has a different meaning. Again we can integrate over both copies of \mathbb{R}^3 in (6.4.49), using

$$\begin{aligned} f_{\xi_1 \xi_2 \xi_3} &:= \int_{-\infty}^{\infty} dD_1 dD_2 dD_3 (D_1^2 - D_2^2)^2 (D_1^2 - D_3^2)^2 (D_2^2 - D_3^2)^2 |D_1 D_2 D_3| e^{-\xi_1 D_1^2 - \xi_2 D_2^2 - \xi_3 D_3^2} \\ &= \frac{24}{\xi_1^5 \xi_2^5 \xi_3^5} \left(2(\xi_1^2 \xi_2^2 + \xi_1^2 \xi_3^2 + \xi_2^2 \xi_3^2)(\xi_1^2 + \xi_2^2 + \xi_3^2 - \xi_1 \xi_2 - \xi_1 \xi_3 - \xi_2 \xi_3) \right. \\ &\quad \left. - 3\xi_1 \xi_2 \xi_3 (\xi_1^3 + \xi_2^3 + \xi_3^3 - \xi_2^2 \xi_3 - \xi_3^2 \xi_1 - \xi_1^2 \xi_2 - \xi_3^2 \xi_2 - \xi_1^2 \xi_3 - \xi_2^2 \xi_1) - 8\xi_1^2 \xi_2^2 \xi_3^2 \right). \end{aligned} \quad (6.4.51)$$

To fix the parameters appearing in the matrices A and A' , we again use expectation values for squared quark masses:

$$\langle D_1^2 \rangle \approx \frac{\int_{\mathbb{R}^3} dD D_1^2 \int_0^{\pi/4} dx \int_0^{\pi/2} dz \sin 2x \sin 2z \left(e^{-\text{Tr}(D^2 U A U^\dagger)} \right) \Big|_{y=0}}{\int_{\mathbb{R}^3} dD \int_0^{\pi/4} dx \int_0^{\pi/2} dz \sin 2x \sin 2z \left(e^{-\text{Tr}(D^2 U A U^\dagger)} \right) \Big|_{y=0}}. \quad (6.4.52)$$

The denominator of (6.4.52) is explicitly

$$I_D := \int_0^{\pi/4} dx \int_0^{\pi/2} dz \sin 2x \sin 2z f_{\xi_1 \xi_2 \xi_3}, \quad (6.4.53)$$

where ξ_1, ξ_2, ξ_3 at $y = 0$ were defined in (6.4.22). Again, we can approximate I_D well by only keeping the terms of leading order in x and z in the trigonometric functions, and approximating $f_{\xi_1 \xi_2 \xi_3}$ by the leading term

$$f_{\xi_1 \xi_2 \xi_3} \approx \frac{24}{\xi_1^5 \xi_2^5 \xi_3^5} \times 2\xi_3^4 \xi_2^2 \approx \frac{48}{(A_1 + A_2 z^2)^5 (A_2 + A_3 x^2)^3 A_3}, \quad (6.4.54)$$

which leads to

$$\begin{aligned}
I_D &\approx \frac{48}{A_3} \int_0^{\pi/4} dx \int_0^{\pi/2} dz \, 4xz (A_1 + A_2 z^2)^{-5} (A_2 + A_3 x^2)^{-3} \\
&\approx \frac{48}{A_3} \int_0^\infty dX \int_0^\infty dZ (A_1 + A_2 Z)^{-5} (A_2 + A_3 X)^{-3} \\
&= \frac{48}{A_3} \cdot \frac{1}{4A_2 A_1^4} \cdot \frac{1}{2A_3 A_2^2} = \frac{6}{A_1^4 A_2^3 A_3^2}.
\end{aligned} \tag{6.4.55}$$

The result is very well reproduced by numerical calculations. Similarly,

$$\begin{aligned}
I_D \langle D_1^2 \rangle &\approx \frac{240}{A_3} \int_0^\infty dX \int_0^\infty dZ (A_1 + A_2 Z)^{-6} (A_2 + A_3 X)^{-3} \\
&= \frac{240}{A_3} \cdot \frac{1}{5A_2 A_1^5} \cdot \frac{1}{2A_3 A_2^2} = \frac{24}{A_1^5 A_2^3 A_3^2},
\end{aligned} \tag{6.4.56}$$

hence

$$\langle D_1^2 \rangle \approx \frac{4}{A_1}. \tag{6.4.57}$$

Redoing the same calculation for D_2 and D_3 gives

$$\langle D_2^2 \rangle \approx \frac{2}{A_2}, \quad \langle D_3^2 \rangle \approx \frac{1}{A_3}. \tag{6.4.58}$$

The relative factors, which are now slightly different from the ones obtained in (6.4.27), have to be taken into account when determining A and A' .

Recall that we used the following central values for the quark masses:

$$\begin{aligned}
(m_u, m_c, m_t) &:= (1.27 \text{ MeV}, 0.619 \text{ GeV}, 171.7 \text{ GeV}) \\
(m_d, m_s, m_b) &:= (2.9 \text{ MeV}, 55 \text{ MeV}, 2.89 \text{ GeV}).
\end{aligned} \tag{6.4.59}$$

We now proceed as before, using numerical calculations to check our analytical approximations. As before, the analytical approximations are better reproduced for the positively charged top, charm and up quarks, which exhibit a more extreme quark mass hierarchy, than for the negatively charged quarks. We use the numerical results to fix the parameters in A and A' to

$$\begin{aligned}
\mu_c^2 &= 2 \left(\frac{m_c}{m_t} \right)^2 \approx 2.60 \times 10^{-5}, \quad \mu_u^2 = 4 \left(\frac{m_u}{m_t} \right)^2 \approx 2.19 \times 10^{-10}, \\
\mu_s^2 &= \left(\frac{m_s}{m_b} \right)^2 \approx 3.62 \times 10^{-4}, \quad \mu_d^2 = 4 \left(\frac{m_d}{m_b} \right)^2 \approx 4.03 \times 10^{-6}.
\end{aligned} \tag{6.4.60}$$

In order to obtain an analytical expression for expectation value of J^2 , we again take the approximation of small z

$$J_{\text{small } z}^2 := \frac{1}{2} \sin^2 x \cos^2 x \sin^2 x' \cos^2 x' \cos^2 z' \sin^4 z' \quad (6.4.61)$$

for our calculations. Within this approximation for J , still taking $f_{\xi_1 \xi_2 \xi_3} \approx 48 \xi_1^{-5} \xi_2^{-3} \xi_3^{-1}$, the numerator of (6.4.49) is the product (using again that only small z contributes)

$$\begin{aligned} & 1152 \times \int_0^{\pi/2} dz \frac{2z}{(A_1 + A_2 z^2)^{5/2}} \times \int_0^{\pi/2} dz' \frac{\sin 2z' \cos^2 z' \sin^4 z'}{(A_1 \cos^2 z' + A_2 \sin^2 z')^5} \\ & \times \int_0^{\pi/4} dx \frac{\sin 2x \sin^2 x \cos^2 x}{(A_2 \cos^2 x + A_3 \sin^2 x)^3 (A_3 \cos^2 x + A_2 \sin^2 x)} \end{aligned} \quad (6.4.62)$$

times the integral in the second line with all quantities replaced by “primed” ones. The first two factors are $\frac{1}{4A_1^4 A_2}$ and $\frac{1}{12(A_1')^2 (A_2')^3}$, respectively; for the other two (which have the same form) we change variables to $X = \cos^2 x$ to obtain

$$\int_{1/2}^1 dX \frac{X(1-X)}{(A_2 X + A_3(1-X))^3 (A_3 X + A_2(1-X))} = \frac{1}{2A_3^3 A_2} \left(1 - \frac{2A_2}{A_3} + O((A_2/A_3)^2) \right); \quad (6.4.63)$$

putting everything together, we obtain

$$\langle J_{\text{small } z}^2 \rangle \approx \frac{1}{6} \frac{A_2 (A_1')^2}{A_3 A_2' A_3'} = \frac{4}{3} \frac{m_s^2 m_u^2 m_d^2}{m_b^4 m_c^2}, \quad (6.4.64)$$

where the numerical factors appearing in the last line come from the different factors chosen in (6.4.60). Just as before, the top quark mass does not appear in this approximation. This compares with the scaling behaviour obtained in the previous section,

$$\langle J_{\text{small } z}^2 \rangle \sim \frac{m_s^2 m_u m_d}{m_b^3 m_c}. \quad (6.4.65)$$

For numerical calculations we use both the simplified expression $J_{\text{small } z}^2$ and the expression for J given in (6.4.19). We find that for the first quantity, the numerically evaluated expectation value, $\langle J_{\text{small } z}^2 \rangle \approx 1.89 \times 10^{-15}$, is about 94% of (6.4.64), and the numerical result for $\langle J^2 \rangle$ (taken at $y = y' = 0$) is

$$\langle J^2 \rangle \approx 2.07 \times 10^{-15}, \quad (6.4.66)$$

which gives

$$\Delta J = \sqrt{\langle J^2 \rangle} \approx 4.55 \times 10^{-8} \quad (6.4.67)$$

which is now almost three orders of magnitude *smaller* than the observed value. Assuming a Gaussian distribution peaked at zero, we now get

$$P(|J| \leq 10^{-7}) \approx 97\%, \quad (6.4.68)$$

When the measure presented here is used, there seems to be extreme fine-tuning in J , but now we would say that one observes unnaturally *large* CP violation! This result may look surprising, given that the maximal value for J is around 0.1 and the observed value just 3×10^{-5} , but it shows how strongly the quark mass hierarchy suppresses large values of J in our distribution.

6.5 Extension to Neutrinos

In the standard model, neutrinos are taken to be massless. There are no Yukawa couplings to the Higgs field, and no possibility of CP violation. However, since the phenomenon of neutrino oscillations has been confirmed experimentally, neutrinos must be assumed to have mass. Curiously, a mixing matrix for leptons, now referred to as the **Maki-Nakagawa-Sakata (MNS) matrix**, was already proposed in 1962 [114]. It appears in a charged current that has the same form as (6.2.7) for baryons, with massive leptons taking the place of up-type quarks and neutrinos taking the place of down-type quarks. With compelling evidence for neutrino oscillations [7] and hence massive neutrinos, such an interaction is now assumed to exist. By diagonalising the leptonic mass terms one will again encounter a mixing matrix M [95].

The mathematical problem is therefore the same as before; the only outstanding issue is the ambiguity in the definition of this matrix due to possible rephasing of the lepton fields. There are two different possible types of mass terms for fermions. In the most commonly used notation, fermion fields are given by complex 4-component (Dirac) spinors, consisting of two 2-component (Weyl) spinors:

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \quad \bar{\psi} := \psi^\dagger \gamma^0 = (\psi_2^\dagger \quad \psi_1^\dagger). \quad (6.5.1)$$

Normally, a mass term contains ψ and its Dirac conjugate $\bar{\psi}$, thus coupling ψ_1 and ψ_2 :

$$\mathcal{L}_{\text{mass}} = m\bar{\psi}\psi = m(\psi_2^\dagger \psi_1 + \psi_1^\dagger \psi_2). \quad (6.5.2)$$

In this case, the four complex components of ψ represent a particle together with its antiparticle, which have the same mass but are otherwise distinct. It is however also

possible to have Majorana mass terms, which are of the form

$$\mathcal{L}_{\text{Majorana}} = m_1 \psi_1^\dagger \psi_1 + m_2 \psi_2^\dagger \psi_2. \quad (6.5.3)$$

Here m_1 and m_2 are not necessarily equal. In the case $m_1 \neq m_2$, the 4-spinor represents two different particles, which are each equal to their own antiparticle. It is then appropriate to take real and imaginary parts and rewrite the theory in terms of two real 4-component (Majorana) spinor fields.^{‡‡}

These fields, being real, have no rephasing freedom, and the matrix relating the flavour and mass eigenbases for neutrinos is in $SO(3)$. Therefore, for an MNS matrix which couples three Majorana neutrinos to the massive leptons, known to be Dirac spinors, there is only the freedom of rephasing on one side, and this matrix would be an element of the single quotient $SU(3)/U(1)^2$. Of course, all intermediate cases are also conceivable, where some neutrinos have Majorana and others have Dirac masses. The most general mixing matrix is hence of the form $U(1)^k \backslash SU(N)/U(1)^{N-1}$, for some k between 0 and $N - 1$.

There are at present no strong constraints on the magnitude of CP violation from observation. The results presented in this chapter may possibly put constraints on it, if there is some experimental information about neutrino masses, or conversely make near-coincident neutrino masses appear more or less likely, if the magnitude of CP violation is known.

6.5.1 Neutrino Mixing Matrix

If we assume that the neutrinos are Majorana spinors, the lepton mixing matrix belongs to $U(1)^2 \backslash SU(3)$, since only phasing of the lepton charge eigenstates $(\nu_e, \nu_\mu, \nu_\tau)$, but not the mass eigenstates (ν_1, ν_2, ν_3) is possible. The former are defined by

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = M \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}. \quad (6.5.4)$$

In other words, $|\nu_\alpha\rangle$ is the neutrino state created in the decay $W^+ \rightarrow l_\alpha^+ + \nu$, where l_α is an antilepton of flavour $\alpha = e, \mu, \tau$ [7]. Such a state is in general a superposition of

^{‡‡}There is of course no need for a theory formulated in terms of Majorana spinors to have an even number of fields, since the Dirac spinor should not be regarded as fundamental. The presentation only started from the Dirac spinor field because it is most commonly encountered in the literature.

mass eigenstates which will have different propagation amplitudes, leading to nonzero probabilities for neutrino oscillation $\nu_\alpha \rightarrow \nu_\beta$.

One conventionally fixes the phases so that M takes the form

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e^{i\alpha_1/2} & 0 & 0 \\ 0 & e^{i\alpha_2/2} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (6.5.5)$$

where the three angles θ_{12} , θ_{13} , and θ_{23} lie in the first quadrant.

The Jarlskog invariant for the neutrino mixing matrix is defined as in (6.2.16) but with (x, y, z, w) now parametrising the MNS matrix M . Note, in particular, that it is independent of the phases α_1 and α_2 .

Experimentally, parameters of the neutrino mixing matrix are not completely known. According to [7],

$$\sin^2(2\theta_{12}) = 0.87 \pm 0.03, \quad 0.92 < \sin^2(2\theta_{23}) \leq 1, \quad \sin^2(2\theta_{13}) < 0.19, \quad (6.5.6)$$

and there is no experimental information about the Dirac angle δ . Thus, we can certainly deduce that there is an upper bound on the Jarlskog invariant for the neutrino mixing matrix, given by

$$|J| < 0.049. \quad (6.5.7)$$

For six different neutrino mass eigenstates, as in the see-saw mechanism, which are assumed to all couple to charged leptons, a general mixing matrix would be an element of $U(1)^5 \backslash SU(6)$, since one would diagonalise a 6×6 Hermitian matrix. The present viewpoint is however that such additional neutrino fields would be “sterile” and not couple to either W or Z bosons [7].

6.5.2 Statistics of J

We have seen that the parameter space for the neutrino mixing matrix is the six-dimensional single quotient $U(1)^2 \backslash SU(3)$, and that the Jarlskog invariant for the neutrino mixing matrix takes the same form (6.2.16) as it does for the CKM matrix. Therefore, all results obtained in section 6.3 apply equally to the case of neutrinos.

For completeness, we quote the results obtained in section 6.3 for a right-invariant measure on the homogeneous space $U(1)^2 \backslash SU(3)$,

$$\langle J^2 \rangle = \frac{1}{720} \approx 1.389 \times 10^{-3}, \quad \langle J^4 \rangle = \frac{1}{201600} \approx 4.960 \times 10^{-5}, \quad (6.5.8)$$

and

$$\Delta J = \frac{1}{12\sqrt{5}} \approx 0.0373. \quad (6.5.9)$$

This can be compared with the experimental bound given in (6.5.7).

One should also repeat the calculations of section 6.4, assuming particular values for the neutrino masses (and including the known masses for the charged leptons). A strong hierarchy in the neutrino masses would then presumably again lead to “naturally” small CP violation from the corresponding mixing matrix. Alternatively, an experimental observation of small CP violation for neutrinos would perhaps be an indication of a mass hierarchy in neutrinos. At present, neither the magnitude of CP violation nor any values of neutrino masses have been measured sufficiently accurately to allow predictions.

6.6 Summary and Outlook

In this chapter, we analysed the problem of finding a natural measure on a space of coupling constants, which in our case was the space of CKM matrices, the double quotient $U(1)^2 \backslash SU(3)/U(1)^2$. We saw that the measure on this double quotient is nonunique, and we analysed several possible choices of measure on the double quotient. One class of measures was given by squashed Kaluza-Klein measures, induced by a Kaluza-Klein reduction of a left-invariant metric on the flag manifold. Alternatively, one could take the unique measure on $SU(3)/U(1)^2$ and simply integrate over the left angles. The measure used by Ozsváth and Schücking seemed not to be very well motivated from a geometric perspective.

When calculating expectation values for J , we found that all of the measures we considered led to rather similar results. In each case, the observed value was about three orders of magnitude below what one would normally expect; the observed value appears to be finely tuned. The same applied to the Ozsváth-Schücking measure, an extremely squashed Kaluza-Klein measure, or a flat measure, which is just the simplest choice and not justified geometrically.

In section 6.4, we adopted the different viewpoint that the CKM matrix should not be viewed as separate from the quark masses, but that it is really the mass matrices which are “chosen” by a yet unknown physical mechanism. We took the observed values for the quark masses as an input and chose the simplest distribution which was able to reproduce these observed values, while inducing a different measure on

the space of CKM matrices. While this is a choice we made, and all results depend on this choice, our measure is the combination of a maximally symmetric measure, invariant under a redefinition of a Hermitian matrix by conjugation by a unitary matrix, and a Gaussian incorporating the observed values of the quark masses. We would have to make additional rather strong assumptions to motivate a different choice of measure that would differ appreciably from this simple construction. It may well be that such assumptions are justified by the underlying mechanism determining the mass matrices, but we do not know of such a mechanism yet. Assuming such a distribution, together with the assumption that the mass matrices can be taken as Hermitian as is possible in the standard model, we found that the observed value of J now appears very natural and not finely tuned at all. In this statistical approach, regarding the Yukawa couplings determining the mass matrices as randomly chosen seems more appropriate than separating quark masses and mixing angles.

In order to test the dependence of this nice result on the assumption of Hermitian mass matrices, we then repeated the calculation under the assumption that there is a left-right symmetry which implies that the quark mass matrices can not in general be taken to be Hermitian. We constructed the analogous probability distribution on the space of 3×3 complex matrices, again fitting four free parameters to expectation values for quark masses that can reproduce the observed values. We saw that using such a probability distribution the conclusion one reaches is rather different: One would now expect J to be about three orders of magnitude smaller than the observed value. Hence, there is a fine-tuning problem; without further assumptions, a fundamental theory leading to a left-right symmetric electroweak sector at low energy should generically be expected to reproduce very weak CP violation. Invoking the principle of Occam's razor, "ENTIA NON SUNT MULTIPLICANDA PRAETER NECES-SITATEM," we would like to conclude that, only looking at possible explanations for the magnitude of CP violation in the electroweak sector, the standard model should be preferred to left-right symmetric extensions such as Pati-Salam: In the latter one needs additional assumptions on the fundamental parameters that resolve the issue of observing "unnaturally large" CP violation, that are not necessary in the standard model, or any extension of it that allows a restriction to Hermitian mass matrices.

In the general context of this thesis, these calculations provide an example of gauge symmetries, described by one or two copies of $SU(2)$, whose presence or absence has a direct effect on measurable quantities, although we have followed a purely statistical

approach rather than proposing dynamical mechanisms that could lead to certain values for these quantities.

Although we have tried to argue that our results are independent of renormalisation group flow since the relevant quantities do not run strongly with energy scale, there is another more subtle issue: The low-energy limit of a left-right symmetric extension with non-Hermitian mass matrices would still be the standard model, where mass matrices can be assumed to be Hermitian, leading us back to the measure we considered previously. It would be desirable to incorporate this dependency of the assumptions one has used to construct the measure on energy scale into the analysis, namely to use a measure which depends on energy scale also. A starting point would be a quantification of “non-Hermiticity” that could then flow from zero at low energies to some non-zero value at high energies. At present these ideas are however somewhat vague, so that we will have to leave them to exploration in future work.

We have argued that our analysis also applies to the case of massive neutrinos, where the predictions will conceivably be tested by future experiments. We have seen that for Majorana spinors, the Maki-Nakagawa-Sakata matrix [114] which appears is naturally an element of the single quotient $U(1)^2 \backslash SU(3)$. Since the right phases do not play any role in neutrino oscillations and the relevant J is independent of these phases, the calculations are identical to the ones presented here, although with the appropriate values of the μ parameters appearing in A and A' . Based on our results, one would expect a strong hierarchy in neutrino masses to lead to suppressed CP violation.

In the seesaw mechanism one adds very heavy right-handed neutrinos, and the most general mixing matrix (making the additional, at present experimentally unjustified, assumption that the heavy neutrinos couple to charged leptons as well) would be an element of $U(1)^5 \backslash SU(6)$. This is naturally a Kähler manifold, and the measure induced by the Kähler metric can be obtained from the analysis in [136]. We leave a detailed treatment of this case, following our approach here, to future work.

Finally, one could analyse the effects of a fourth generation of quarks on CP violation by repeating the calculations for 4×4 Hermitian matrices. If this generalisation spoils the agreement with the observed J , one might obtain interesting lower bounds on the masses of a hypothetical extra generation of quarks.

Summary and Outlook

*“La recherche de la vérité doit être le but de notre activité;
c’est la seule fin qui soit digne d’elle.” [140]*

Henri Poincaré

In this thesis we have investigated several examples within general relativity and particle physics where symmetries, and more precisely the geometry of symmetry groups and homogeneous spaces, played a fundamental role. In each of these cases, the study of geometric aspects of the underlying symmetries led to physical predictions, providing illuminating examples for the interplay between mathematical structures and the physical world that modern theoretical physics is based on.

The first part consisting of chapters 2 to 4 was concerned with deformed and constrained symmetries in relativity.

In chapter 2, we studied exact solutions of general relativity in four dimensions from a group-theoretic viewpoint, focussing on their global symmetries (isometries). The isometry groups of these spacetimes act simply-transitively; the spacetimes are homogeneous spaces G/H , where H was either trivial or one-dimensional. We applied the theory of deformations of Lie groups and algebras to these isometry groups, finding that some solutions are related by deformation of their isometry groups, and that these algebraic properties are partly related to physical properties; we have at least found indications that the Kaigorodov-Ozsváth solution contains closed timelike curves (CTCs) just as the Petrov solution which it is a deformation of. This work, which focussed on a few concrete relatively straightforward examples, leaves much scope for extension. First, one could go to higher dimensions, since in four dimensions all homogeneous solutions are more or less known. Second, one could investigate what the generic relationship is between algebraic properties (related to isometry groups) and physical properties (such as causal behaviour); one might be able to relate a

classification of group deformations to a classification of physical properties of certain classes of exact solutions. We leave this to further work.

In chapter 3, we considered a formulation of general relativity as a gauge theory of the de Sitter group, shifting our focus to local symmetries. This formulation, an extension of the MacDowell-Mansouri formulation given by Stelle and West, replaces the perhaps more usual Poincaré group by the de Sitter group, and provides another example of group deformation. While the viewpoint on general relativity as some kind of gauge theory is not new, we argued how Einstein-Cartan-Stelle-West theory may provide a “deformation” of general relativity analogous to the framework of deformed special relativity (DSR), and focussed on Minkowski space as an explicit example, thereby uncovering close relations between two apparently disconnected branches of the general relativity literature. We showed how the apparent “noncommutativity of spacetime” given by noncommuting translations on the internal (“momentum”) de Sitter space is manifest in torsion of the connection. We gave estimates for the torsion one would get in the simplest situation, and discussed worrying physical consequences if electromagnetic fields couple to torsion, such as breaking of charge conservation. While this means that the framework we discussed is presumably not viable as a physical theory, it provided an interesting alternative interpretation of DSR.

Our study was an application of Cartan geometry which is a generalisation of (pseudo-)Riemannian geometry in that the “tangent” space is replaced by an arbitrary homogeneous space G/H . It would be worth formulating other examples of such a situation in physics in this mathematical language. One example is the traditional formulation of loop quantum gravity (LQG) where the $\mathfrak{so}(3,1)$ -valued connection is reduced to an $\mathfrak{so}(3)$ -valued connection by partial gauge fixing. A recent study [38] seems to generalise this gauge choice by essentially working in a Cartan geometric construction, but the precise relation is rather unclear to us.

In chapter 4 we studied yet another reformulation of general relativity, as a topological BF theory with constraints. Instead of deforming a symmetry, we were constraining it; the large symmetry on the phase space of the topological theory which implies the (local) equivalence of all solutions to the equations of motion could be constrained to be only the gauge symmetry of general relativity in four dimensions, where not all solutions are locally equivalent. Such a formulation, which is popular in covariant approaches to quantising gravity, has so far mainly been given in terms of quadratic constraints, which suffer from ambiguities in their solution whose relevance

in a quantum theory is not entirely clear. It is therefore an improvement if, as we have shown, an alternative formulation in terms of linear constraints can be found. This formulation seems closely related to discrete constraints on representations that are used in spin foam models, but requires in its corresponding classical formulation the introduction of a new “4D closure” constraint that has not been studied very much so far. An obvious extension of our work would be to study the effects of the presence of such a new constraint on spin foam models for quantum gravity. Linear constraints, whose solutions are formally given by projections, are also more closely related to the group field theory approach to quantum gravity, where constraints are usually enforced as projections [121].

The second part consisting of chapters 5 and 6 was a geometric study of gauge symmetries in the electroweak model and their role in CP violation.

In chapter 5 we introduced a class of metrics on the group $SU(3)$, the homogeneous space $SU(3)/U(1)^2$ and the double quotient space $U(1)^2 \backslash SU(3)/U(1)^2$. While for the first two cases the $SU(3)$ left group action provided a clear criterion (left-invariance) for preferred classes of metrics, the situation for the double quotient was less clear and several possible choices were introduced. We also computed bi-invariant metrics (with respect to the natural group actions) on the space of Hermitian and on the space of complex 3×3 matrices.

In chapter 6 we saw how CP violation in the electroweak sector is determined by the Jarlskog invariant J which can be viewed as a function on $U(1)^2 \backslash SU(3)/U(1)^2$, the space of Cabibbo-Kobayashi-Maskawa (CKM) matrices. We tried to give a mathematically precise answer to the question whether the observed magnitude of J , which appears to be rather small, should be considered natural or fine-tuned. Since finding this answer is equivalent to specifying a natural measure on the double quotient space, we analysed several possible choices of the measure induced by metrics presented in the previous chapter. While it seems not entirely clear which of these metrics should be preferred, we found that they all suggest fine-tuning in the magnitude of J . We then reformulated the problem by considering the space of quark mass matrices instead, using a measure which incorporates the observed quark masses and modifies the induced distribution on the space of CKM matrices. The effect of this modified distribution was the following: If the mass matrices can be taken to be Hermitian as in the standard model, we found that the input of observed quark masses is suf-

ficient to reproduce the observed magnitude of CP violation to a remarkably high accuracy. If on the other hand one takes general complex mass matrices, one finds an “inverse” fine-tuning problem; the magnitude of J should be expected to be much smaller than is observed. This distinction is physically relevant since in models with extended left-right symmetry, such as the Pati-Salam model, one has to consider general complex mass matrices. Our results, which rely on assuming the simplest and geometrically best motivated probability distributions, seem therefore to suggest that the standard model should be preferred by Occam’s razor: One does not need additional fine-tuning assumptions to explain the magnitude of CP violation. Hence, these calculations provide another example of how the modification of symmetries (the absence or presence of a left-right symmetry) leads to different physical (in this case statistical) predictions.

Similar calculations could be performed for the case of four quark generations, where the mass of a fourth quark would have an effect on the probability distributions we have considered. Another important extension of our results is to the case of neutrinos, where the formalism for CP violation is rather similar and the mathematics essentially the same. Our results would suggest that a large mass hierarchy in neutrinos (together with the known hierarchy in lepton masses) makes highly suppressed CP violation likely, whereas on the other hand large CP violation might suggest that near-coincident neutrino masses appear more likely. Such predictions can obviously be tested in experiments such as those at LHC.

References (with pages of appearance)

- [1] EMANUELE ALESCI and CARLO ROVELLI. *Complete LQG propagator: Difficulties with the Barrett-Crane vertex*. Physical Review D **76** (2007) 104012. [arXiv:0708.0883](#). / 90
- [2] SERGEI ALEXANDROV. *Simplicity and closure constraints in spin foam models of gravity*. Physical Review D **78** (2008) 044033. [arXiv:0802.3389](#). / 103, 104
- [3] SERGEI ALEXANDROV and ETERA R. LIVINE. *$SU(2)$ loop quantum gravity seen from covariant theory*. Physical Review D **67** (2003) 044009. [arXiv:gr-qc/0209105](#). / 103, 104
- [4] G. AMELINO-CAMELIA. *Testable scenario for relativity with minimum length*. Physics Letters B **510** (2001) 255–263. [arXiv:hep-th/0012238](#). / 9, 58
- [5] G. AMELINO-CAMELIA, JOHN R. ELLIS, N. E. MAVROMATOS, DIMITRI V. NANOPOULOS, and SUBIR SARKAR. *Potential Sensitivity of Gamma-Ray Burster Observations to Wave Dispersion in Vacuo*. Nature **393** (1998) 763–765. [arXiv:astro-ph/9712103](#). / 58
- [6] GIOVANNI AMELINO-CAMELIA, LEE SMOLIN, and ARTEM STARODUBTSEV. *Quantum symmetry, the cosmological constant and Planck-scale phenomenology*. Classical and Quantum Gravity **21** (2004) 3095–3110. [arXiv:hep-th/0306134](#). / 9, 74
- [7] CLAUDE AMSLER ET AL. (Particle Data Group). *Review of Particle Physics*. Phys. Lett. B **667** (2008) 1–1340. / 128, 135, 144, 151, 152, 153
- [8] ASIMINA ARVANITAKI, SAVAS DIMOPOULOS, ANDREW A. GERACI, JASON HOGAN, and MARK KASEVICH. *How to Test Atom and Neutron Neutrality with Atom Interferometry*. Physical Review Letters **100** (2008) 120407. [arXiv:0711.4636](#). / 77

- [9] ANDREAS ARVANITOYEORGOS. *An Introduction to Lie Groups and the Geometry of Homogeneous Spaces*. American Mathematical Society, 2003. / 12, 13, 14, 110, 111
- [10] ABHAY ASHTEKAR. *New variables for classical and quantum gravity*. Physical Review Letters **57** (1986) 2244–2247. / 83
- [11] SAINT AUGUSTINE. *The Confessions (Oxford World’s Classics)*. Oxford Paperbacks, 2008. / 80
- [12] HENRI BACRY and JEAN-MARC LÉVY-LEBLOND. *Possible Kinematics*. Journal of Mathematical Physics **9** (1968) 1605–1614. / 19, 23, 24, 56
- [13] JOHN C. BAEZ. *An introduction to spin foam models of BF theory and quantum gravity*. Lect. Notes Phys. **543** (2000) 25–94. [arXiv:gr-qc/9905087](#). / 81
- [14] JOHN C. BAEZ and JOHN W. BARRETT. *The Quantum Tetrahedron in 3 and 4 Dimensions* [arXiv:gr-qc/9903060](#). / 89
- [15] VIJAY BALASUBRAMANIAN and PER KRAUS. *A Stress Tensor for Anti-de Sitter Gravity*. Communications in Mathematical Physics **208** (1999) 413–428. [arXiv:hep-th/9902121](#). / 29
- [16] ARISTIDE BARATIN and DANIELE ORITI. *Group field theory with non-commutative metric variables* [arXiv:1002.4723](#). / 86, 90, 103, 104, 105
- [17] JOHN W. BARRETT and LOUIS CRANE. *Relativistic spin networks and quantum gravity*. Journal of Mathematical Physics **39** (1998) 3296–3302. [arXiv:gr-qc/9709028](#). / 87
- [18] JOHN W. BARRETT, RICHARD J. DOWDALL, WINSTON J. FAIRBAIRN, HENRIQUE GOMES, and FRANK HELLMANN. *Asymptotic analysis of the Engle-Pereira-Rovelli-Livine four-simplex amplitude*. Journal of Mathematical Physics **50** (2009) 112504. [arXiv:0902.1170](#). / 90
- [19] JOHN W. BARRETT and CHRISTOPHER M. STEELE. *Asymptotics of relativistic spin networks*. Classical and Quantum Gravity **20** (2003) 1341–1361. [arXiv:gr-qc/0209023](#). / 90

-
- [20] THOMAS BAYES. *An Essay towards solving a Problem in the Doctrine of Chances*. Philosophical Transactions (1683-1775) **53** (1763) 370–418. / 121
- [21] I. M. BENN, T. DERELI, and R. W. TUCKER. *Gauge field interactions in spaces with arbitrary torsion*. Physics Letters B **96** (1980) 100–104. / 75
- [22] Z. BERN, J. J. M. CARRASCO, L. J. DIXON, H. JOHANSSON, and R. ROIBAN. *Ultraviolet Behavior of $N=8$ Supergravity at Four Loops*. Physical Review Letters **103** (2009) 081301. 0905.2326. / 9
- [23] EUGENIO BIANCHI and CARLO ROVELLI. *Why all these prejudices against a constant?* arXiv:1002.3966. / 121
- [24] I. I. BIGI and A. I. SANDA. *CP Violation*. Cambridge University Press, 2000. / 135
- [25] W. B. BONNOR. *A source for Petrov’s homogeneous vacuum space-time*. Physics Letters A **75** (1979) 25–26. / 27
- [26] W. B. BONNOR. *The rigidly rotating relativistic dust cylinder*. Journal of Physics A Mathematical General **13** (1980) 2121–2132. / 28
- [27] W. B. BONNOR, N. O. SANTOS, and M. A. H. MACCALLUM. *An exterior for the Gödel spacetime*. Classical and Quantum Gravity **15** (1998) 357–366. arXiv:gr-qc/9711011. / 49
- [28] E. BUFFENOIR, M. HENNEAUX, K. NOUI, and P. ROCHE. *Hamiltonian analysis of Plebanski theory*. Classical and Quantum Gravity **21** (2004) 5203–5220. arXiv:gr-qc/0404041. / 85, 103
- [29] S. CACCIATORI, V. GORINI, and A. KAMENSHCHIK. *Special relativity in the 21st century*. Annalen der Physik **17** (2008) 728–768. arXiv:0807.3009. / 59
- [30] C. CAPRINI and P. G. FERREIRA. *Constraints on the electrical charge asymmetry of the universe*. Journal of Cosmology and Astro-Particle Physics **2** (2005) 6. arXiv:hep-ph/0310066. / 77, 78
- [31] STEVEN CARLIP. *Quantum Gravity in 2+1 Dimensions*. Cambridge University Press, 2003. / 81, 83

- [32] E. CARTAN. *Les groupes d'holonomie des espaces généralisés*. Acta Mathematica **48** (1926) 1–42. / 60
- [33] B. CARTER. *Domains of stationary communications in space-time*. General Relativity and Gravitation **9** (1978) 437–450. / 41
- [34] BRANDON CARTER. *Global Structure of the Kerr Family of Gravitational Fields*. Physical Review **174** (1968) 1559–1571. / 41
- [35] M. CASELLE, A. D'ADDA, and L. MAGNEA. *Regge calculus as a local theory of the Poincaré group*. Physics Letters B **232** (1989) 457–461. / 85
- [36] SHAO-HSUAN CHIU, T. K. KUO, TAE-HUN LEE, and CHI XIONG. *Mass matrices and their renormalization*. Physical Review D **79** (2009) 013012. arXiv:0811.1806. / 122
- [37] C. CHRYSSOMALAKOS and E. OKON. *Generalized Quantum Relativistic Kinematics: a Stability Point of View*. International Journal of Modern Physics D **13** (2004) 2003–2034. arXiv:hep-th/0410212. / 19, 25, 74
- [38] FRANCESCO CIANFRANI and GIOVANNI MONTANI. *Towards Loop Quantum Gravity without the Time Gauge*. Physical Review Letters **102** (2009) 091301. arXiv:0811.1916. / 158
- [39] ANDREW G. COHEN and SHELDON L. GLASHOW. *Very Special Relativity*. Physical Review Letters **97** (2006) 021601. arXiv:hep-ph/0601236. / 19
- [40] FLORIAN CONRADY and LAURENT FREIDEL. *Semiclassical limit of 4-dimensional spin foam models*. Physical Review D **78** (2008) 104023. arXiv:0809.2280. / 85, 90, 91
- [41] ROBERT COQUEREAUX and GILLES ESPOSITO-FARESE. *Right-invariant metrics on the Lie group $SU(3)$ and the Gell-Mann-Okubo formula*. Journal of Mathematical Physics **32** (1991) 826–831. / 109
- [42] M. CVETIČ, H. LÜ, and C. N. POPE. *Space-times of boosted p-branes, and CFT in infinite-momentum frame*. Nuclear Physics B **545** (1999) 309–339. arXiv:hep-th/9810123. / 42

- [43] R. DEBEVER. *Local tetrad in general relativity*. In: *IV Centenario Della Nascita di Galileo Galilei, 1564-1964*. Firenze, 1965 p. 192. / 26
- [44] L. DEFRISE. *Groupes d'isotropie et groupes de stabilité conforme dans les espaces lorentziens*. Ph.D. thesis, Université Libre de Bruxelles (1969). / 48
- [45] R. DEPIETRI and L. FREIDEL. *$so(4)$ Plebanski action and relativistic spin-foam model*. Classical and Quantum Gravity **16** (1999) 2187–2196. arXiv:gr-qc/9804071. / 86, 87
- [46] PAUL ADRIEN MAURICE DIRAC. *Lectures on quantum mechanics*. Yeshiva Univ., Phys. Dept., 1964. / 10
- [47] BIANCA DITTRICH and JAMES P. RYAN. *Phase space descriptions for simplicial 4d geometries* arXiv:0807.2806. / 84, 85, 98, 101, 104
- [48] MICHAEL R. DOUGLAS and NIKITA A. NEKRASOV. *Noncommutative field theory*. Reviews of Modern Physics **73** (2001) 977–1029. arXiv:hep-th/0106048. / 58
- [49] JONATHAN ENGLE, ROBERTO PEREIRA, and CARLO ROVELLI. *Flipped spinfoam vertex and loop gravity*. Nuclear Physics B **798** (2008) 251–290. arXiv:0708.1236. / 86, 90, 97, 98, 103
- [50] EUCLID and SIR THOMAS HEATH (ED.). *The Thirteen Books of Euclid's Elements (3 vol.)*. Dover Publications; 2nd edition, 1956. / 14
- [51] L.E. EVTUSHIK, YU.G. LUMISTE, N.M. OSTIANU, and A.P. SHIROKOV. *Differential-geometric structures on manifolds*. J. Sov. Math. **14** (1979) 1573–1719. / 60
- [52] JOSÉ FIGUEROA-O'FARRILL. *Deformations of M-theory Killing superalgebras*. Classical and Quantum Gravity **24** (2007) 5257–5276. arXiv:0706.2600. / 20
- [53] JOSÉ M. FIGUEROA-O'FARRILL. *Deformations of the Galilean algebra*. Journal of Mathematical Physics **30** (1989) 2735–2739. / 23
- [54] ELKE FLATAU. *Albert Einstein als wissenschaftlicher Autor* Max Planck Institute for the History of Science Preprint 293 (2005). / 50

- [55] P. H. FRAMPTON and C. JARLSKOG. *Systematics of quark mass matrices in the standard electroweak model*. Physics Letters B **154** (1985) 421–424. / 126
- [56] L. FREIDEL, K. KRASNOV, and R. PUZIO. *BF description of higher-dimensional gravity theories*. Adv. Theor. Math. Phys. **3** (1999) 1289–1324. [arXiv:hep-th/9901069](#). / 81
- [57] LAURENT FREIDEL and KIRILL KRASNOV. *A new spin foam model for 4D gravity*. Classical and Quantum Gravity **25** (2008) 125018. [arXiv:0708.1595](#). / 86, 90, 97, 98, 103
- [58] LAURENT FREIDEL and ARTEM STARODUBTSEV. *Quantum gravity in terms of topological observables* [arXiv:hep-th/0501191](#). / 74
- [59] ALFONSO GARCÍA-PARRADO and JOSÉ M. M. SENOVILLA. *Causal structures and causal boundaries*. Classical and Quantum Gravity **22** (2005) R1–R84. [arXiv:gr-qc/0501069](#). / 41
- [60] MURRAY GELL-MANN. *The Eightfold Way: A Theory of strong interaction symmetry* California Institute of Technology Synchrotron Lab. preprint CTSL-20 (March 1961). / 107
- [61] G. W. GIBBONS. *Talk given at The Very Early Universe 25 Years On, Cambridge, 20th December 2007*. / 121
- [62] G. W. GIBBONS, JOAQUIM GOMIS, and C. N. POPE. *General very special relativity is Finsler geometry*. Physical Review D **76** (2007) 081701. [arXiv:0707.2174](#). / 20, 21, 23
- [63] G. W. GIBBONS, JOAQUIM GOMIS, and C. N. POPE. *Deforming the Maxwell-Sim Algebra* 0910.3220. / 20
- [64] G. W. GIBBONS and C. E. PATRICOT. *Newton Hooke spacetimes, Hpp-waves and the cosmological constant*. Classical and Quantum Gravity **20** (2003) 5225–5239. [arXiv:hep-th/0308200](#). / 23
- [65] GARY W. GIBBONS and STEFFEN GIELEN. *The Petrov and Kaigorodov Ozsváth solutions: spacetime as a group manifold*. Classical and Quantum Gravity **25** (2008) 165009. [arXiv:0802.4082](#). / 2, 15

- [66] GARY W. GIBBONS and STEFFEN GIELEN. *Deformed general relativity and torsion*. Classical and Quantum Gravity **26** (2009) 135005. [arXiv:0902.2001](#). / 2, 15
- [67] GARY W. GIBBONS, STEFFEN GIELEN, C. N. POPE, and NEIL TUROK. *Measures on mixing angles*. Physical Review D **79** (2009) 013009. [arXiv:0810.4813](#). / 2, 16, 131
- [68] STEFFEN GIELEN. *CP violation makes left-right symmetric extensions with non-Hermitian mass matrices appear unnatural*. Physical Review D **81** (2010) 076003. [arXiv:1002.0342](#). / 2, 16
- [69] STEFFEN GIELEN and DANIELE ORITI. *Classical general relativity as BF-Plebanski theory with linear constraints*. Classical and Quantum Gravity **27** (2010) 185017. [arXiv:1004.5371](#). / 2, 15
- [70] FLORIAN GIRELLI, ETERA R. LIVINE, and DANIELE ORITI. *Deformed special relativity as an effective flat limit of quantum gravity*. Nuclear Physics B **708** (2005) 411–433. [arXiv:gr-qc/0406100](#). / 74
- [71] DOMENICO GIULINI. *A Euclidean Bianchi model based on $S^3/(D_8^*)$* . Journal of Geometry and Physics **20** (1996) 149–159. [arXiv:gr-qc/9508040](#). / 115
- [72] S. GLASHOW. *Partial-symmetries of weak interactions*. Nuclear Physics **22** (1961) 579–588. / 9, 107
- [73] YU. A. GOL’FAND. *On the introduction of an “elementary length” in the relativistic theory of elementary particles*. Soviet Physics JETP **10** (1960) 356–360. / 57
- [74] YU. A. GOL’FAND. *Quantum field theory in constant curvature p-space*. Soviet Physics JETP **16** (1963) 184–191. / 57, 58
- [75] G. GRIGNANI and G. NARDELLI. *Gravity and the Poincaré group*. Physical Review D **45** (1992) 2719–2731. / 62
- [76] S. S. GUBSER, I. R. KLEBANOV, and A. M. POLYAKOV. *Gauge theory correlators from non-critical string theory*. Physics Letters B **428** (1998) 105–114. [arXiv:hep-th/9802109](#). / 38

- [77] S. W. HAWKING. *Chronology protection conjecture*. Physical Review D **46** (1992) 603–611. / 26
- [78] F. W. HEHL, J. D. MCCREA, E. W. MIELKE, and Y. NE’EMAN. *Metric-affine gauge theory of gravity: field equations, Noether identities, world spinors, and breaking of dilation invariance*. Physics Reports **258** (1995) 1–171. arXiv:gr-qc/9402012. / 78
- [79] MIROSLAV HEJNA. *Symmetric Affine Theories and Nonlinear Einstein-Proca System* arXiv:gr-qc/0611118. / 77, 78
- [80] M. HENNINGSON and K. SKENDERIS. *The holographic Weyl anomaly*. Journal of High Energy Physics **7** (1998) 23. arXiv:hep-th/9806087. / 29
- [81] CHRISTOPHER P. HERZOG, MUKUND RANGAMANI, and SIMON F. ROSS. *Heating up Galilean holography*. Journal of High Energy Physics **11** (2008) 80. arXiv:0807.1099. / 20
- [82] DAVID HILBERT. *Die Grundlagen der Physik*. Mathematische Annalen **92** (1924) 1–32. / 84, 85
- [83] SÖREN HOLST. *Barbero’s Hamiltonian derived from a generalized Hilbert-Palatini action*. Physical Review D **53** (1996) 5966–5969. arXiv:gr-qc/9511026. / 84, 85
- [84] GARY T. HOROWITZ. *Exactly soluble diffeomorphism invariant theories*. Communications in Mathematical Physics **125** (1989) 417–437. / 84
- [85] F. HOYLE. *A New Model for the Expanding Universe*. Monthly Notices of the Royal Astronomical Society **108** (1948) 372. / 77
- [86] F. HOYLE. *On the Possible Consequences of a Variability of the Elementary Charge*. Royal Society of London Proceedings Series A **257** (1960) 431–442. / 77
- [87] PATRICK HUET and ERIC SATHER. *Electroweak baryogenesis and standard model CP violation*. Physical Review D **51** (1995) 379–394. arXiv:hep-ph/9404302. / 121

- [88] E. INONU and E. P. WIGNER. *On the Contraction of Groups and Their Representations*. Proceedings of the National Academy of Science **39** (1953) 510–524. / 21, 24
- [89] C. JARLSKOG. *Commutator of the quark mass matrices in the standard electroweak model and a measure of maximal CP nonconservation*. Physical Review Letters **55** (1985) 1039–1042. / 126, 127, 136
- [90] C. JARLSKOG. *Flavor projection operators and applications to CP violation with any number of families*. Physical Review D **36** (1987) 2128–2136. / 128
- [91] C. JARLSKOG (Editor). *CP Violation*. World Scientific, 1989. / 16, 124, 127
- [92] V. G. KADYSHEVSKII. *On the theory of quantization of space-time*. Soviet Physics JETP **14** (1962) 1340–1346. / 57, 58
- [93] V. R. KAIGORODOV. *Einstein Spaces of Maximum Mobility*. Soviet Physics Doklady **7** (1963) 893–895. / 28, 42
- [94] T. W. B. KIBBLE. *Lorentz Invariance and the Gravitational Field*. Journal of Mathematical Physics **2** (1961) 212–221. / 54
- [95] S. F. KING. *Neutrino mass*. Contemporary Physics **48** (2007) 195–211. arXiv:0712.1750. / 151
- [96] FELIX KLEIN. *Vergleichende Betrachtungen über neuere geometrische Forschungen*. Mathematische Annalen **43** (1893) 63–100. / 7, 51
- [97] FELIX KLEIN. *Vorlesungen über nicht-Euklidische Geometrie*. Chelsea Publishing Co./American Mathematical Society, 2000. / 14
- [98] M. KOBAYASHI and T. MASKAWA. *CP-Violation in the Renormalizable Theory of Weak Interaction*. Progress of Theoretical Physics **49** (1973) 652–657. / 122, 127
- [99] SHOSHICHI KOBAYASHI and KATSUMI NOMIZU. *Foundations of Differential Geometry, Volume I*. Interscience Publ., 1963. / 67
- [100] JERZY KOWALSKI-GLIKMAN. *De Sitter space as an arena for doubly special relativity*. Phys. Lett. **B547** (2002) 291–296. arXiv:hep-th/0207279. / 58, 59

- [101] KIRILL KRASNOV. *Plebanski Formulation of General Relativity: A Practical Introduction* arXiv:0904.0423. / 84
- [102] CLAUS LÄMMERZAHN, ALFREDO MACÍAS, and HOLGER MÜLLER. *Lorentz invariance violation and charge (non)conservation: A general theoretical frame for extensions of the Maxwell equations*. Physical Review D **71** (2005) 025007. arXiv:gr-qc/0501048. / 77
- [103] KORNEIL LANZOS. *Über eine stationäre Kosmologie im Sinne der Einsteinschen Gravitationstheorie*. Zeitschrift für Physik **21** (1924) 73–110. / 27
- [104] MONIQUE LEVY-NAHAS. *Deformation and Contraction of Lie Algebras*. Journal of Mathematical Physics **8** (1967) 1211–1222. / 21
- [105] E. R. LIVINE and S. SPEZIALE. *Solving the simplicity constraints for spinfoam quantum gravity*. Europhysics Letters **81** (2008) 50004. arXiv:0708.1915. / 86, 89, 97, 100
- [106] ETERA R. LIVINE. *Projected spin networks for Lorentz connection: linking spin foams and loop gravity*. Classical and Quantum Gravity **19** (2002) 5525–5541. arXiv:gr-qc/0207084. / 103, 104
- [107] RICHARD E. LIVINE and DANIELE ORITI. *Barrett-Crane spin foam model from generalized BF-type action for gravity*. Physical Review D **65** (2002) 044025. arXiv:gr-qc/0104043. / 90
- [108] R. A. LYTTLETON and H. BONDI. *On the Physical Consequences of a General Excess of Charge*. Royal Society of London Proceedings Series A **252** (1959) 313–333. / 77
- [109] M. A. H. MACCALLUM and N. O. SANTOS. *Stationary and static cylindrically symmetric Einstein spaces of the Lewis form*. Classical and Quantum Gravity **15** (1998) 1627–1636. arXiv:gr-qc/9712028. / 29
- [110] S. W. MACDOWELL and F. MANSOURI. *Unified geometric theory of gravity and supergravity*. Physical Review Letters **38** (1977) 739–742. / 54, 60, 63
- [111] J. MADORE. *Noncommutative Geometry for Pedestrians*. In: P. G. BERGMANN, V. DE SABBATA, and N. J. GOLDBERG (Editors), *Classical and Quantum Nonlocality*. 2000 p. 111. arXiv:gr-qc/9906059. / 74

- [112] JOÃO MAGUEIJO and LEE SMOLIN. *Lorentz Invariance with an Invariant Energy Scale*. Physical Review Letters **88** (2002) 190403. [arXiv:hep-th/0112090](#). / 59, 60
- [113] S. MAJID and H. RUEGG. *Bicrossproduct structure of κ -Poincare group and non-commutative geometry*. Physics Letters B **334** (1994) 348–354. [arXiv:hep-th/9405107](#). / 58
- [114] ZIRO MAKI, MASAMI NAKAGAWA, and SHOICHI SAKATA. *Remarks on the Unified Model of Elementary Particles*. Progress of Theoretical Physics **28** (1962) 870–880. / 111, 151, 156
- [115] M. MARINELLI and G. MORPURGO. *The electric neutrality of matter: A summary*. Physics Letters B **137** (1984) 439–442. / 77
- [116] R. VILELA MENDES. *Deformations, stable theories and fundamental constants*. Journal of Physics A: Mathematical and General **27** (1994) 8091–8104. / 18, 19, 24, 25
- [117] SIMONE MERCURI. *A possible topological interpretation of the Barbero-Immirzi parameter* [arXiv:0903.2270](#). / 85
- [118] ALBERT ABRAHAM MICHELSON and EDWARD MORLEY. *On the Relative Motion of the Earth and the Luminiferous Ether*. American Journal of Science **34** (1887) 333–345. / 8
- [119] J. W. MOFFAT. *Quantum gravity momentum representation and maximum invariant energy* [arXiv:gr-qc/0401117](#). / 60
- [120] M. NAKAHARA. *Geometry, topology and physics, 2nd Edition*. Taylor & Francis, 2003. / 54
- [121] DANIELE ORITI. *The group field theory approach to quantum gravity* [arXiv:gr-qc/0607032](#). / 84, 86, 159
- [122] DANIELE ORITI. *Approaches to Quantum Gravity: Toward a New Understanding of Space, Time and Matter*. Cambridge University Press, 2009. / 84
- [123] DANIELE ORITI. *Group field theory and simplicial quantum gravity* [arXiv:0902.3903](#). / 86, 103

- [124] S. ORITO and M. YOSHIMURA. *Can the universe be charged?* Physical Review Letters **54** (1985) 2457–2460. / 77
- [125] ISTVÁN OZSVÁTH. *All homogeneous solution of Einstein's vacuum field equations with a non vanishing cosmological term.* In: W. RINDLER and A. TRAUTMAN (Editors), *Gravitation and Geometry*. Bibliopolis, Napoli, 1987 p. 309. / 28
- [126] ISTVÁN OZSVÁTH. *Working with Engelbert.* In: A. HARVEY (Editor), *On Einstein's Path: essays in honor of Engelbert Schucking*. 1999 p. 339. / 112
- [127] A. PAPAPETROU. *Champs gravitationnels stationnaires à symétrie axiale.* Annales de l'institut Henri Poincaré (A) Physique théorique **4** (1966) 83–105. / 29
- [128] J. C. PATI and A. SALAM. *Lepton number as the fourth "color".* Physical Review D **10** (1974) 275–289. / 126
- [129] C. PATRICOT. *Kaigorodov spaces and their Penrose limits.* Classical and Quantum Gravity **20** (2003) 2087–2102. arXiv:hep-th/0302073. / 42
- [130] WOLFGANG PAULI. *I. Das Jahr 1930.* In: *Wissenschaftlicher Briefwechsel mit Bohr, Einstein, Heisenberg u.a.: Band 2: 1930-1939*. Springer, Berlin, 1979 . / 55
- [131] PETER PELDAN. *Actions for gravity, with generalizations: A Review.* Class. Quant. Grav. **11** (1994) 1087–1132. arXiv:gr-qc/9305011. / 83
- [132] R. PENROSE. *The road to reality: a complete guide to the laws of the universe.* Vintage, 2004. / 107
- [133] ALEJANDRO PEREZ. *Spin foam models for quantum gravity.* Classical and Quantum Gravity **20** (2003) R43–R104. arXiv:gr-qc/0301113. / 84, 86, 87
- [134] A. Z. PETROV. *Gravitational field geometry as the geometry of automorphisms.* In: *Recent Developments in General Relativity*. Pergamon, Oxford, 1962 p. 379. / 26
- [135] R. J. PETTI. *Translational spacetime symmetries in gravitational theories.* Classical and Quantum Gravity **23** (2006) 737–751. / 67, 68

- [136] R. F. PICKEN. *The Duistermaat-Heckman integration formula on flag manifolds*. Journal of Mathematical Physics **31** (1990) 616–638. / 111, 112, 156
- [137] MAX PLANCK. *Über irreversible Strahlungsvorgänge*. Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften zu Berlin **15** (1899) 440–480. / 55
- [138] JERZY F. PLEBAŃSKI. *On the separation of Einsteinian substructures*. Journal of Mathematical Physics **18** (1977) 2511–2520. / 81, 84
- [139] JIŘÍ PODOLSKÝ. *Interpretation of the Siklos solutions as exact gravitational waves in the anti-de Sitter universe*. Classical and Quantum Gravity **15** (1998) 719–733. arXiv:gr-qc/9801052. / 42, 44
- [140] HENRI POINCARÉ. *La valeur de la science*. Flammarion, 1905. / 157
- [141] GEORG RAFFELT. *Pulsar bound on the photon electric charge reexamined*. Physical Review D **50** (1994) 7729–7730. arXiv:hep-ph/9409461. / 77
- [142] MICHAEL P. REISENBERGER. *A lattice worldsheet sum for 4-d Euclidean general relativity* arXiv:gr-qc/9711052. / 87
- [143] MICHAEL P. REISENBERGER. *Classical Euclidean general relativity from ‘left-handed area = right-handed area’*. Classical and Quantum Gravity **16** (1999) 1357–1371. arXiv:gr-qc/9804061. / 86, 87, 94
- [144] R. ROSENFELD and J. L. ROSNER. *Hierarchy and anarchy in quark mass matrices, or can hierarchy tolerate anarchy?* Physics Letters B **516** (2001) 408–414. arXiv:hep-ph/0106335. / 142
- [145] CARLO ROVELLI. *Quantum Gravity*. Cambridge University Press, 2004. / 84
- [146] ABDUS SALAM. *Weak and Electromagnetic Interactions*. In: SVARTHOLM (Editor), *Elementary Particle Theory, Proceedings Of The Nobel Symposium Held 1968 At Lerum, Sweden*. 1968 pp. 367–377. / 9, 107
- [147] FREDERIC PAUL SCHULLER. *Dirac-Born-Infeld kinematics, maximal acceleration and almost product manifolds*. Ph.D. thesis, University of Cambridge (2003). / 56

- [148] R. W. SHARPE. *Differential Geometry: Cartan's Generalization of Klein's Erlangen Program*. Springer-Verlag, New York, 1997. / 51, 61, 68
- [149] S. T. C. SIKLOS. *Lobatchevski plane gravitational waves*. In: M. A. H. MACCALLUM (Editor), *Galaxies, axisymmetric systems and relativity; Essays presented to W.B. Bonnor on his 65th birthday*. Cambridge University Press, 1985 pp. 247–274. / 42, 48, 49
- [150] LEE SMOLIN and ARTEM STARODUBTSEV. *General relativity with a topological phase: an action principle* [arXiv:hep-th/0311163](#). / 84
- [151] HARTLAND S. SNYDER. *Quantized Space-Time*. Physical Review **71** (1947) 38–41. / 25, 52, 55
- [152] ALAN R. STEIF. *Quantum stress tensor in the three-dimensional black hole*. Physical Review D**49** (1994) 585–589. [arXiv:gr-qc/9308032](#). / 20
- [153] K. S. STELLE and P. C. WEST. *Spontaneously broken de Sitter symmetry and the gravitational holonomy group*. Physical Review D**21** (1980) 1466–1488. / 51, 60, 61, 64, 65, 66
- [154] HANS STEPHANI, DIETRICH KRAMER, MALCOLM MACCALLUM, CORNELIUS HOENSELAERS, and EDUARD HERLT. *Exact Solutions of Einstein's Field Equations, 2nd Edition*. Cambridge University Press, 2003. / 14, 25, 26, 28, 42, 48
- [155] W. J. VAN STOCKUM. *The gravitational field of a distribution of particles rotating around an axis of symmetry*. Proc. Roy. Soc. Edinburgh **57** (1937) 135–154. / 27
- [156] G. 'T HOOFT and M. VELTMAN. *Regularization and renormalization of gauge fields*. Nuclear Physics B **44** (1972) 189–213. / 9
- [157] FRANK J. TIPLER. *Rotating cylinders and the possibility of global causality violation*. Physical Review D**9** (1974) 2203–2206. / 27, 37, 40, 41, 49
- [158] M. ULANOVSKII. *On global properties of certain homogeneous exact solutions of Einstein field equations*. Journal of Mathematical Physics **49** (2008) 052502. / 40, 41

-
- [159] RYOYU UTIYAMA. *Invariant Theoretical Interpretation of Interaction*. Physical Review **101** (1956) 1597–1607. / 54
- [160] H. WAELBROECK and J. A. ZAPATA. *A Hamiltonian lattice formulation of topological gravity*. Classical and Quantum Gravity **11** (1994) 989–998. arXiv:gr-qc/9311035. / 85
- [161] STEVEN WEINBERG. *A Model of Leptons*. Physical Review Letters **19** (1967) 1264–1266. / 9, 107
- [162] P. WEST. *E_{11} and M theory*. Classical and Quantum Gravity **18** (2001) 4443–4460. arXiv:hep-th/0104081. / 20
- [163] HERMANN WEYL. *Raum. Zeit. Materie: Vorlesungen über allgemeine Relativitätstheorie, 5., umgearb. Aufl.* Springer, Berlin, 1923. / 54, 96
- [164] HERMANN WEYL. *Gruppentheorie und Quantenmechanik*. S. Hirzel in Leipzig, 1928. / 7, 11
- [165] HERMANN WEYL. *A Remark on the Coupling of Gravitation and Electron*. Phys. Rev. **77** (1950) 699–701. / 54
- [166] DEREK K. WISE. *MacDowell-Mansouri gravity and Cartan geometry* arXiv:gr-qc/0611154. / 61, 63, 65, 67, 68
- [167] DEREK K. WISE. *Symmetric space Cartan connections and gravity in three and four dimensions* arXiv:0904.1738. / 61, 70
- [168] EDWARD WITTEN. *2 + 1 dimensional gravity as an exactly soluble system*. Nuclear Physics B **311** (1988) 46–78. / 81
- [169] LINCOLN WOLFENSTEIN. *Parametrization of the Kobayashi-Maskawa Matrix*. Physical Review Letters **51** (1983) 1945–1947. / 134
- [170] ZHI-ZHONG XING, HE ZHANG, and SHUN ZHOU. *Updated values of running quark and lepton masses*. Physical Review **D77** (2008) 113016. arXiv:0712.1419. / 122, 142

List of Figures

2.1	Possible contractions and deformations of kinematical groups.	24
3.1	Cartan geometry extends both Klein geometry and (pseudo-)Riemannian geometry.	51
3.2	Rolling an internal de Sitter space along a closed path in Minkowski space.	67
4.1	Lagrangians and Hamiltonians for general relativity in four dimensions.	83
4.2	A 4-simplex is generated by five points.	88
4.3	Dual picture of a 4-simplex.	99
6.1	Probability distribution for $\log J $ using the $SU(3)$ -invariant flag measure.	133
6.2	Probability distribution for $\log J $ using the Kaluza-Klein measure.	133
6.3	Probability distribution for $\log J $ using a uniform distribution.	134

List of Tables

2.1	Infinitesimal deformations of the Petrov Killing algebra.	31
2.2	Infinitesimal deformations of the Kaigorodov Killing algebra.	45
6.1	Analytical and numerical results for integrals of interest.	146

Index

- \mathbb{Z}_2 grading, 70
- 4-simplex, 84
- AdS/CFT, 20, 29, 49
- anti-de Sitter space, 39, 40
- Ashtekar new variables, 83
- Bayes' theorem, 121
- BF theory, 80
- bi-invariant metric
 - for complex matrices, 118
 - for Hermitian matrices, 115
 - on $SU(3)$, 110
- Bianchi classification, 26, 43
- Cartan geometry, 51
- chirality, 58
- CKM matrix, 126, 128
- closed timelike curves, 20, 27, 40
- cohomology, 23
- constraints, 10, 84, 86
 - first class, 10, 102
 - Gauss, 89, 102
 - simplicity, 86
- cosmic microwave background (CMB), 78
- cosmological constant problem, 63, 121
- creation field, 77
- cylinder, rotating, 27, 36
- de Sitter space, 56, 59, 70
- deformed special relativity (DSR), 58
- development, 65, 71, 75
- dihedral group \mathfrak{S}_3 , 115
- dilatation, 38
- discrete symmetries, 11
- dominant energy condition, 35
- double quotient, 107, 154
- effective potential, 37
- Einstein-Cartan theory, 53
- Einstein-Cartan-Stelle-West theory, 64
- Einstein-Hilbert action, 52, 63, 81, 85
- Einstein-Proca theory, 77
- electroweak model, 9, 107
- equivalence principle, 50
- Erlangen Programme, 7, 51
- Euclid's fifth postulate, 14
- Euclidean group, 8
- Euler angles, 108, 128
- Euler-Lagrange equations, 39
- exact sequence, 68
- experiment, 77, 153
- fine-tuning, 120, 124
- flag manifold, 111, 131
- gauge, 96
- gauge theory, 8, 54
- gauge transformation, 54, 62, 92
- Gell-Mann matrices, 108
- geodesic completeness, 38

-
- gravitational radiation, 44, 48
 - Higgs, 125
 - homogeneous bundles, 60
 - homogeneous space, 13, 129
 - Jacobi identity, 21, 45
 - Jarlskog invariant, 122, 127, 128
 - Kaigorodov solution, 42
 - Kaigorodov-Ozsváth solution, 28
 - Kaluza-Klein reduction, 113, 114, 130
 - Killing form, 110
 - Killing vector, 26, 28, 42, 47
 - kinematical group, 23, 56
 - left-invariant metric, 13, 31, 34, 36, 39, 44, 46, 109
 - left-invariant one-form, 30
 - left-invariant vector field, 12
 - Lewis form, 29
 - LHC, 160
 - Lie algebra, 12, 54, 62
 - Lie bracket, 12
 - Lie group, 11, 122
 - Lie group/algebra deformation, 8, 19, 21, 31, 45, 56
 - Lie's theorems, 12
 - M-theory, 20
 - MacDowell-Mansouri formulation, 53, 63
 - mass hierarchy, 123, 154
 - Maurer-Cartan form, 12, 30, 33–35, 39, 44, 46, 109, 130
 - Maurer-Cartan relations, 13, 23
 - minimal coupling, 75
 - MNS matrix, 151
 - Moyal bracket, 24
 - Occam's razor, 155
 - orthonormal frame, 48, 51, 65, 81, 96
 - Pati-Salam model, 123, 126, 155
 - Petrov classification, 26, 28, 42
 - Petrov solution, 26
 - physical units, 73
 - point particle, 37
 - pp-wave, 42
 - priors, 121
 - quantum chromodynamics, 107
 - quantum gravity, 9, 55, 60, 74, 81, 85
 - loop, 84
 - spin foam approach, 81
 - quark mass matrix, 125
 - reductive geometry, 14, 44, 69, 111
 - renormalisation, 9, 122, 142, 156
 - Schrödinger spacetime, 20
 - Schwarzschild, 76
 - Snyder's non-commuting coordinates, 25, 52, 56, 66
 - special relativity, 58
 - spinor, 151
 - Majorana, 152
 - stability of theories, 18
 - Stokes' theorem, 89
 - string theory, 9, 42, 55
 - structure constants, 13, 22, 70
 - supersymmetry, 9
 - symmetric affine theory, 78
 - symmetric space, 70
 - symmetry breaking, 138

- symplectic structure, 10
- time-orientable, 27, 29
- topological field theory, 10, 80
- torsion, 52, 66, 71, 72, 75
- totally vicious, 41
- transformation group, 7
- unimodular, 110
- unitarity triangle, 127
- weak energy condition, 34, 36
- Witten, 81
- Wolfenstein parametrisation, 134
- zero section, 67